PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07D 207/34, 233/90, A61K 31/415, C07D 403/14, C12Q 1/68

A1

(11) International Publication Number: WO 98/37066

(43) International Publication Date: 27 August 1998 (27.08.98)

(21) International Application Number: PCT/US98/01006

(22) International Filing Date: 21 January 1998 (21.01.98)

(30) Priority Data:

PCT7US97/03332	20 February 1997 (20.02.97)	WO
(34) Countries for which the regional or international application was filed:		US et al.
60/043,444	8 April 1997 (08.04.97)	US
60/042,022	16 April 1997 (16.04.97)	US
08/837,524	21 April 1997 (21.04.97)	US
08/853,522	8 May 1997 (08.05.97)	US
PCT/US97/12722	21 July 1997 (21.07.97)	WO
(34) Countries for which the regional or international application was filed:		US et al.

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	08/853,522 (CIP)
Filed on	8 May 1997 (08.05.97)
US	08/837,524 (CIP)
Filed on	21 April 1997 (21.04.97)
US	08/607,078 (CIP)
Filed on	26 February 1996 (26.02.96)
US	60/042,022 (CIP)
Filed on	16 April 1997 (16.04.97)
US	60/043,444 (CIP)
Filed on	8 April 1997 (08.04.97)

(71) Applicant *(for all designated States except US)*: CALIFORNIA INSTITUTE OF TECHNOLOGY [US/US]; Pasadena, CA 91125 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BAIRD, Eldon, E. [US/US]; 255 S. Madison Avenue #5, Pasadena, CA 91101 (US). DERVAN, Peter, B. [US/US]; 1235 St. Albans Road, San Marino, CA 91108 (US).

(74) Agent: McDONNELL, John, J.; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Chigago, IL 60606 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: IMPROVED POLYAMIDES FOR BINDING IN THE MINOR GROOVE OF DOUBLE STRANDED DNA

(57) Abstract

The invention encompasses improved polyamides for binding to specific nucleotide sequences in the minor groove of double stranded DNA. The 3–hydroxy–N–methylpyrrole/N–methylpyrrole carboxamide pair specifically recognizes the T.A base pair, while the N–methylpyrrole/3–hydroxy–N–methylpyrrole pair recognizes A.T nucleotide pairs. Similarly, an N–methylimidizole/N–methylpyrrole carboxamide pair specifically recognizes the G.C nucleotide pair, and the N–methylpyrrole/N–methylimidizole carboxamide pair recognizes the C.G nucleotide pair.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IMPROVED POLYAMIDES FOR BINDING IN THE MINOR GROOVE OF DOUBLE STRANDED DNA

The U.S. Government has certain rights in this invention pursuant to Grant Nos. GM 26453, 27681 and 47530 awarded by the National Institute of Health.

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of PCT/US97/03332 filed February 20, 1997, Serial No. 08/853,522 filed May 8, 1997 and PCT/US 97/12722 filed July 21, 1997 which are continuation-in-part applications of Serial No. 08/837,524, filed April 21, 1997, Serial No. 08/607,078, filed February 26, 1996, provisional application Serial No. 60/042,022, filed April 16, 1997 and provisional application Serial No. 60/043,444, filed April 8, 1997.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to polyamides which bind to predetermined sequences in the minor groove of double stranded DNA.

Description of the Related Art

The design of synthetic ligands that read the information stored in the DNA double helix has been a long standing goal of chemistry. Cell-permeable small molecules which target predetermined DNA sequences are useful for the regulation of gene-expression. Oligodeoxynucleotides that recognize the major groove of double-helical DNA via triple-helix formation bind to a broad range of sequences with high affinity and specificity. Although oligonucleotides and their analogs have been shown to interfere with gene expression, the triple helix approach is limited to purine tracks and suffers from poor cellular uptake. The development of pairing rules for minor groove binding polyamides derived from N-methylpyrrole (Py) and N-methylimidazole (Im) amino acids provides another code to control sequence specificity. An Im/Py pair distinguishes G•C from C•G and both of these from A•T or T•A base pairs. Wade, W.S., Mrksich, M. & Dervan, P.B. describes the design of peptides that bind in the minor groove of DNA at 5'-(A,T)G(A,T)C(A,T)-3' sequences by a dimeric side-by-side motif. *J. Am. Chem. Soc.* **114**, 8783-8794 (1992); Mrksich, M. *et al.* describes antiparallel side-by-side motif for sequence specific-recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamidenetropsin. *Proc. Natl. Acad. Sci. USA* **89**, 7586-7590 (1992); Trauger, J.W., Baird, E. E. Dervan, P.B. describes the recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **382**, 559-561 (1996). A

Py/Py pair specifies A•T from G•C but does not distinguish A•T from T•A. Pelton, J.G. & Wemmer, D.E. describes the structural characterization of a 2-1 distamycin A-d(CGCAAATTTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA* **86**, 5723-5727 (1989); White, S., Baird, E. E. & Dervan, P.B. Describes the effects of the A•T/T•A degeneracy of pyrrole-imidazole polyamide recognition in the minor groove of DNA. *Biochemistry* **35**, 6147-6152 (1996); White, S., Baird, E. E. & Dervan, P. B. describes the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem. & Biol.* **4**, 569-578 (1997); White, S., Baird, E. E. & Dervan, P.B. describes the 5'-3' N-C orientation preference for polyamide binding in the minor groove. In order to break this degeneracy, a new aromatic amino acid, 3-hydroxy-N-methylpyrrole (Hp) incorporated into a polyamide and paired opposite Py, has been found to discriminate A•T from T•A. The replacement of a single hydrogen atom on the pyrrole with a hydroxy group in a Hp/Py pair regulates affinity and specificity of a polyamide by an order of magnitude. Utilizing Hp together with Py and Im in polyamides to form four aromatic amino acid pairs (Im/Py, Py/Im, Hp/Py, and Py/Hp) provides a code to distinguish all four Watson-Crick base pairs in the minor groove of DNA.

SUMMARY OF THE INVENTION

The invention encompasses improved polyamides for binding to the minor groove of double stranded ("duplex") DNA. The polyamides are in the form of a hairpin comprising two groups of at least three consecutive carboxamide residues, the two groups covalently linked by an aliphatic amino acid residue, preferably γ-aminobutyric acid or 2,4 diaminobutyric acid, the consecutive carboxamide residues of the first group pairing in an antiparallel manner with the consecutive carboxamide residues of the second group in the minor groove of double stranded DNA. The improvement relates to the inclusion of a binding pair of Hp/Py carboxamides in the polyamide to bind to a T•A base pair in the minor groove of double stranded DNA or Py/Hp carboxamide binding pair in the polyamide to bind to an A•T base pair in the minor groove of double stranded DNA. The improved polyamides have at least three consecutive carboxamide pairs for binding to at least three DNA base pairs in the minor groove of a duplex DNA sequence that has at least one A•T or T•A DNA base pair, the improvement comprising selecting a Hp/Py carboxamide pair to correspond to a T•A base pair in the minor groove or a Py/Hp carboxamide pair to bind to an A•T DNA base pair in the minor groove. Preferably the binding of the carboxamide pairs to the DNA base pairs modulates the expression of a gene.

In one preferred embodiment, the polyamide includes at least four consecutive carboxamide pairs for binding to at least four base pairs in a duplex DNA sequence. In another preferred embodiment, the polyamide includes at least five consecutive carboxamide pairs for binding to at least five base pairs in a duplex DNA sequence. In yet another preferred

embodiment, the polyamide includes at least six consecutive carboxamide pairs for binding to at least six base pairs in a duplex DNA sequence. In one preferred embodiment, the improved polyamides have four carboxamide binding pairs that will distinguish A•T, T•A, C•G and G•C base pairs in the minor groove of a duplex DNA sequence. The duplex DNA sequence can be a regulatory sequence, such as a promoter sequence or an enhancer sequence, or a gene sequence, such as a coding sequence or a non-coding sequence. Preferably, the duplex DNA sequence is a promoter sequence.

The preparation and the use of polyamides for binding in the minor groove of double stranded DNA are extensively described in the art. This invention is an improvement of the existing technology that uses 3-hydroxy-N-methylpyrrole to provide carboxamide binding pairs for DNA binding polyamides.

The invention encompasses polyamides having γ-aminobutyric acid or a substituted γ-aminobutyric acid to form a hairpin with a member of each carboxamide pairing on each side of it. Preferably the substituted γ-aminobutyric acid is a chiral substituted γ-aminobutyric acid such as (R)-2,4-diaminobutyric acid. In addition, the polyamides may contain an aliphatic amino acid residue, preferably a β-alanine residue, in place of a non-Hp carboxamide. The β-alanine residue is represented in formulas as β. The β-alanine residue becomes a member of a carboxamide binding pair. The invention further includes the substitution as a β•β binding pair for non-Hp containing binding pair. Thus, binding pairs in addition to the Hp/Py and Py/Hp are Im/β, β/Im, Py/β, β/Py, and β/β.

The polyamides of the invention can have additional moieties attached covalently to the polyamide. Preferably the additional moieties are attached as substituents at the amino terminus of the polyamide, the carboxy terminus of the polyamide, or at a chiral (R)-2,4-diaminobutyric acid residue. Suitable additional moieties include a detectable labeling group such as a dye, biotin or a hapten. Other suitable additional moieties are DNA reactive moieties that provide for sequence specific cleavage of the duplex DNA.

Brief Description of the Drawings

Figure 1 illustrates the structure of polyamide 1, 2, and 3.
Figure 2 illustrates the pairing of polyamides to DNA base pairs.
Figure 3 illustrates the DNase footprint titration of compounds 2 and 3.
Figure 4 illustrates a list of the structures of representative Hp containing polyamides.
Figure 5 illustrates the synthesis of a protected Hp monomer for solid phase synthesis.
Figure 6 illustrates the solid phase synthesis of polyamide 2.
Figure 7 illustrates the 1H-NMR characterization of polyamide 2.

Figure 8 illustrates the Mass spectral characterization of polyamide **2**.
Figure.9 illustrates 1H-NMR characterization of synthesis purity.
Figure 10 illustrates DNaseI footprint titration experiment.
Figure 11 illustrates the synthesis of bifunctional conjugate of polyamide **2**.
Figure 12 illustrates affinity cleaving evidence for oriented hairpin formation.
Figure 13 illustrates increased sequence specificity of Hp/Py containing polyamides.
Figure 14 illustrates 8-ring hairpin polyamides which target 5'-WGTNNW-3' sites.
Figure 15 illustrates 8-ring hairpin polyamides which target 5'-WGANNW-3' sites.
Figure 16 illustrates 8-ring hairpin polyamides which target 5'-WGGNNW-3' sites.
Figure 17 illustrates 8-ring hairpin polyamides which target 5'-WGCNNW-3' sites.

DETAILED DESCRIPTION OF THE INVENTION

Within this application, unless otherwise stated, definitions of the terms and illustration of the techniques of this application may be found in any of several well-known references such as: Sambrook, J., *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989); Goeddel, D., *ed.*, *Gene Expression Technology*, *Methods in Enzymology*, **185**, Academic Press, San Diego, CA (1991); "Guide to Protein Purification" in Deutshcer, M.P., ed., *Methods in Enzymology*, Academic Press, San Diego, CA (1989); Innis, et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press, San Diego, CA (1990); Freshney, R.I., *Culture of Animal Cells: A Manual of Basic Technique, 2nd Ed.*, Alan Liss, Inc. New York, NY (1987); Murray, E.J., *ed.*, *Gene Transfer and Expression Protocols*, pp. 109-128, The Humana Press Inc., Clifton, NJ and Lewin, B., *Genes VI*, Oxford University Press, New York (1997).

For the purposes of this application, a *promoter* is a regulatory sequence of DNA that is involved in the binding of RNA polymerase to initiate transcription of a gene. A *gene* is a segment of DNA involved in producing a peptide, polypeptide or protein, including the coding region, non-coding regions preceding ("leader") and following ("trailer") the coding region, as well as intervening non-coding sequences ("introns") between individual coding segments ("exons"). Coding refers to the representation of amino acids, start and stop signals in a three base "triplet" code. Promoters are often upstream (" '5 to") the transcription initiation site of the corresponding gene. Other regulatory sequences of DNA in addition to promoters are known, including sequences involved with the binding of transcription factors, including response elements that are the DNA sequences bound by inducible factors. Enhancers comprise

yet another group of regulatory sequences of DNA that can increase the utilization of promoters, and can function in either orientation (5'-3' or 3'-5') and in any location (upstream or downstream) relative to the promoter. Preferably, the regulatory sequence has a positive activity, i.e., binding of an endogeneous ligand (e.g. a transcription factor) to the regulatory sequence increases transcription, thereby resulting in increased expression of the corresponding target gene. In such a case, interference with transcription by binding a polyamide to a regulatory sequence would reduce or abolish expression of a gene.

The promoter may also include or be adjacent to a regulatory sequence known in the art as a *silencer*. A silencer sequence generally has a negative regulatory effect on expression of the gene. In such a case, expression of a gene may be increased directly by using a polyamide to prevent binding of a factor to a silencer regulatory sequence or indirectly, by using a polyamide to block transcription of a factor to a silencer regulatory sequence.

It is to be understood that the polyamides of this invention bind to double stranded DNA in a sequence specific manner. The function of a segment of DNA of a given sequence, such as 5'-TATAAA-3', depends on its position relative to other functional regions in the DNA sequence. In this case, if the sequence 5'-TATAAA-3' on the coding strand of DNA is positioned about 30 base pairs upstream of the transcription start site, the sequence forms part of the promoter region (Lewin, *Genes VI,* pp. 831-835). On the other hand, if the sequence 5'-TATAAA-3' is downstream of the transcription start site in a coding region and in proper register with the reading frame, the sequence encodes the tyrosyl and lysyl amino acid residues (Lewin, *Genes VI,* pp. 213-215).

While not being held to one hypothesis, it is believed that the binding of the polyamides of this invention modulate gene expression by altering the binding of DNA binding proteins, such as RNA polymerase, transcription factors, TBF, TFIIIB and other proteins. The effect on gene expression of polyamide binding to a segment of double stranded DNA is believed to be related to the function, e.g., promoter, of that segment of DNA.

It is to be understood by one skilled in the art that the improved polyamides of the present invention may bind to any of the above-described DNA sequences or any other sequence having a desired effect upon expression of a gene. In addition, U.S. Patent No.

5,578,444 describes numerous promoter targeting sequences from which base pair sequences for targeting an improved polyamide of the present invention may be identified.

It is generally understood by those skilled in the art that the basic structure of DNA in a living cell includes both *major* and a *minor groove*. For the purposes of describing the present invention, the *minor groove* is the narrow groove of DNA as illustrated in common molecular biology references such as Lewin, B., *Genes VI*, Oxford University Press, New York (1997).

To affect gene expression in a cell, which may include causing an increase or a decrease in gene expression, a effective quantity of one or more polyamide is contacted with the cell and internalized by the cell. The cell may be contacted *in vivo* or *in vitro*. Effective extracellular concentrations of polyamides that can modulate gene expression range from about 10 nanomolar to about 1 micromolar. Gottesfeld, J.M., *et al.*, *Nature* 387 202-205 (1997). To determine effective amounts and concentrations of polyamides *in vitro*, a suitable number of cells is plated on tissue culture plates and various quantities of one or more polyamide are added to separate wells. Gene expression following exposure to a polyamide can be monitored in the cells or medium by detecting the amount of the protein gene product present as determined by various techniques utilizing specific antibodies, including ELISA and western blot. Alternatively, gene expression following exposure to a polyamide can be monitored by detecting the amount of messenger RNA present as determined by various techniques, including northern blot and RT-PCR.

Similarly, to determine effective amounts and concentrations of polyamides for *in vivo* administration, a sample of body tissue or fluid, such as plasma, blood, urine, cerebrospinal fluid, saliva, or biopsy of skin, muscle, liver, brain or other appropriate tissue source is analyzed. Gene expression following exposure to a polyamide can be monitored by detecting the amount of the protein gene product present as determined by various techniques utilizing specific antibodies, including ELISA and western blot. Alternatively, gene expression following exposure to a polyamide can be monitored by the detecting the amount of messenger RNA present as determined by various techniques, including northern blot and RT-PCR.

The polyamides of this invention may be formulated into diagnostic and therapeutic compositions for *in vivo* or *in vitro* use. Representative methods of formulation may be found

in *Remington: The Science and Practice of Pharmacy*, 19th ed., Mack Publishing Co., Easton, PA (1995).

For *in vivo* use, the polyamides may be incorporated into a physiologically acceptable pharmaceutical composition that is administered to a patient in need of treatment or an animal for medical or research purposes. The polyamide composition comprises pharmaceutically acceptable carriers, excipients, adjuvants, stabilizers, and vehicles. The composition may be in solid, liquid, gel, or aerosol form. The polyamide composition of the present invention may be administered in various dosage forms orally, parentally, by inhalation spray, rectally, or topically. The term parenteral as used herein includes, subcutaneous, intravenous, intramuscular, intrasternal, infusion techniques or intraperitoneally.

The selection of the precise concentration, composition, and delivery regimen is influenced by, *inter alia*, the specific pharmacological properties of the particular selected compound, the intended use, the nature and severity of the condition being treated or diagnosed, the age, weight, gender, physical condition and mental acuity of the intended recipient as well as the route of administration. Such considerations are within the purview of the skilled artisan. Thus, the dosage regimen may vary widely, but can be determined routinely using standard methods.

Polyamides of the present invention are also useful for detecting the presence of double stranded DNA of a specific sequence for diagnostic or preparative purposes. The sample containing the double stranded DNA can be contacted by polyamide linked to a solid substrate, thereby isolating DNA comprising a desired sequence. Alternatively, polyamides linked to a suitable detectable marker, such as biotin, a hapten, a radioisotope or a dye molecule, can be contacted by a sample containing double stranded DNA.

The design of bifunctional sequence specific DNA binding molecules requires the integration of two separate entities: recognition and functional activity. Polyamides that specifically bind with subnanomolar affinity to the minor groove of a predetermined sequence of double stranded DNA are linked to a functional molecule, providing the corresponding bifunctional conjugates useful in molecular biology, genomic sequencing, and human medicine. Polyamides of this invention can be conjugated to a variety of functional molecules, which can be independently chosen from but is not limited to arylboronic acids, biotins, polyhistidines

comprised from about 2 to 8 amino acids, haptens to which an antibody binds, solid phase supports, oligodeoxynucleotides, N-ethylnitrosourea, fluorescein, bromoacetamide, iodoacetamide, DL-α-lipoic acid, acridine, captothesin, pyrene, mitomycin, texas red, anthracene, anthrinilic acid, avidin, DAPI, isosulfan blue, malachite green, psoralen, ethyl red, 4-(psoraen-8-yloxy)-butyrate, tartaric acid, (+)-α-tocopheral, psoralen, EDTA, methidium, acridine, Ni(II)•Gly-Gly-His, TO, Dansyl, pyrene, N-bromoacetamide, and gold particles. Such bifunctional polyamides are useful for DNA affinity capture, covalent DNA modification, oxidative DNA cleavage, DNA photocleavage. Such bifunctional polyamides are useful for DNA detection by providing a polyamide linked to a detectable label. Detailed instructions for synthesis of such bifunctional polyamides can be found in copending U.S. provisional application 60/043,444, the teachings of which are incorporated by reference.

DNA complexed to a labeled polyamide can then be determined using the appropriate detection system as is well known to one skilled in the art. For example, DNA associated with a polyamide linked to biotin can be detected by a streptavidin / alkaline phosphatase system.

The present invention also describes a diagnostic system, preferably in kit form, for assaying for the presence of the double stranded DNA sequence bound by the polyamide of this invention in a body sample, such brain tissue, cell suspensions or tissue sections, or body fluid samples such as CSF, blood, plasma or serum, where it is desirable to detect the presence, and preferably the amount, of the double stranded DNA sequence bound by the polyamide in the sample according to the diagnostic methods described herein.

The diagnostic system includes, in an amount sufficient to perform at least one assay, a specific polyamide as a separately packaged reagent. Instructions for use of the packaged reagent(s) are also typically included. As used herein, the term "package" refers to a solid matrix or material such as glass, plastic (e.g., polyethylene, polypropylene or polycarbonate), paper, foil and the like capable of holding within fixed limits a polyamide of the present invention. Thus, for example, a package can be a glass vial used to contain milligram quantities of a contemplated polyamide or it can be a microliter plate well to which microgram quantities of a contemplated polypamide have been operatively affixed, i.e., linked so as to be capable of being bound by the target DNA sequence. "Instructions for use" typically include a tangible expression describing the reagent concentration or at least one assay method parameter such as the relative amounts of reagent and sample to be admixed, maintenance time

periods for reagent or sample admixtures, temperature, buffer conditions and the like. A diagnostic system of the present invention preferably also includes a detectable label and a detecting or indicating means capable of signaling the binding of the contemplated polyamide of the present invention to the target DNA sequence. As noted above, numerous detectable labels, such as biotin, and detecting or indicating means, such as enzyme-linked (direct or indirect) streptavidin, are well known in the art.

Figure 1 shows representative structures of polyamides. ImImPyPy-γ-ImPyPyPy-β-Dp (**1**), ImImPyPy-γ-ImHpPyPy-β-Dp (**2**), and ImImHpPy-γ-ImPyPyPy-β-Dp (**3**). (Hp = 3-hydroxy-N-methylpyrrole, Im = N-methylimidazole, Py = N-methylpyrrole, β = β-alanine, γ = γ-aminobutyric acid, Dp = Dimethylaminopropylamide). Polyamides were synthesized by solid phase methods using Boc-protected 3-methoxypyrrole, imidazole, and pyrrole aromatic amino acids, cleaved from the support by aminolysis, deprotected with sodium thiophenoxide, and purified by reversed phase HPLC. Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. The identity and purity of the polyamides were verified by ^{1}H NMR, analytical HPLC, and matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS-monoisotopic): **1** 1223.6 (1223.6 calculated), **2** 1239.6 (1239.6 calculated); **3** 1239.6 (1239.6 calculated).

Figure 2 illustrates binding models for polyamides **1-3** in complex with 5'-TGGTCA-3' and 5'-TGGACA-3' (A•T and T•A in fourth position highlighted). Filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole, the curved line connecting the polyamide subunits represents γ-aminobutyric acid, the diamond represents β-alanine, and the + represents the positively charged dimethylaminopropylamide tail group.

Figure 3 shows quantitative DNase I footprint titration experiments with polyamides **2** and **3** on the 3' ^{32}P labeled 250-bp pJK6 *Eco*RI/*Pvu*II restriction fragment. Lane 1, intact DNA; lanes 2-11 DNase I digestion products in the presence of 100, 50, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1 nM polyamide, respectively; lane 12, DNase I digestion products in the absence of polyamide; lane 13, adenine-specific chemical sequencing. Iverson, B. L. & Dervan, P. B. describes an adenine-specific DNA chemical sequencing reaction. *Methods Enzymol.* **15**, 7823-7830 (1987).

All reactions were done in a total volume of 400 μL. A polyamide stock solution or H_2O was added to an assay buffer containing radiolabeled restriction fragment, with the final solution conditions of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0. Solutions were allowed to equilibrate for 4-12 h at 22 °C before initiation of footprinting reactions. Footprinting reactions, separation of cleavage products, and data analysis were carried out as described. White, S., Baird, E. E. & Dervan, P. B. Effects of the A•T/T•A degeneracy of pyrrole-imidazole polyamide recognition in the minor groove of DNA. *Biochemistry 35*, 6147-6152 (1996).

Figure 4 shows the structure and equilibrium dissociation constant for numerous compounds of the present invention. Polyamides are shown in complex with their respective match site. Filled and unfilled circles represent imidazole (Im) and pyrrole (Py) rings, respectively; circles containing an H represent 3-hydroxypyrrole (Hp), the curved line connecting the polyamide subunits represents γ-aminobutyric acid (γ), the diamond represents β-alanine (β), and the + represents the positively charged dimethylaminopropylamide tail group (Dp). The equilibrium dissociation constants are the average values obtained from three DNase I footprint titration experiments. The standard deviation for each set is less than 15% of the reported number. Assays were carried out in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22°C.

Figure 5 shows the synthetic scheme for 3-O-methyl-N-Boc protected pyrrole-2-carboxylate. The hydroxypyrrole monoester can be prepared in 0.5 kg quantity using published procedures on enlarged scale.

Figure 6 shows the solid phase synthetic scheme for ImImPyPy-γ-ImHpPyPy-β-Dp starting from commercially available Boc-β-Pam-Resin: (i) 80% TFA/DCM, 0.4 M PhSH; (ii) Boc-Py-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Py-OBt, DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) Boc-3-OMe-Py-OH, HBTU, DMF, DIEA; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Boc-Im-OH, DCC, HOBt; (ix) 80% TFA/DCM, 0.4 M PhSH; (x) Boc-γ-aminobutyric acid, DIEA, DMF; (xi) 80% TFA/DCM, 0.4 M PhSH; (xii) Boc-Py-OBt, DIEA, DMF; (xiii) 80% TFA/DCM, 0.4 M PhSH; (xiv) Boc-Py-OBt, DMF, DIEA; (xv) 80% TFA/DCM, 0.4 M PhSH; (vxi) Boc-Im-OH, DCC, HOBt (xvii) 80% TFA/DCM, 0.4 M PhSH; (xviii) imidazole-2-carboxylic acid, HBTU, DIEA; (xviv) dimethylaminopropylamine, 55 °C, 18h. Purification by reversed phase HPLC provides ImImPyPy-γ-ImOpPyPy-β-Dp. (Op = 3-methoxypyrrole). Treatment of the 3-methyoxypyrrole polyamide with thiophenol, NaH, DMF, at 100 °C for 120 min provides polyamide **2** after reverse phase HPLC purification.

Figure 7 shows the aromatic region from 7-11 ppm for the 1H-NMR spectrum determined at 300 MHz for ImImPyPy-γ-ImOpPyPy-β-Dp and ImImPyPy-γ-ImHpPyPy-β-Dp. This region of the spectrum may be used to determine compound identity and purity.

Figure 8 shows the MALDI-TOF mass spectrum determined in positive ion mode with a monoisotopic detector for the polyamides for ImImPyPy-γ-ImOpPyPy-β-Dp and ImImPyPy-γ-ImHpPyPy-β-Dp. This spectrum may be used to determine compound identity and purity.

Figure 9 shows the methyl group region from 3.5-4.0 ppm for the 1H-NMR spectrum determined at 300 MHz for ImPyPy-γ-OpPyPy-β-Dp and ImPyPy-γ-HpPyPy-β-Dp. This region of the spectrum may be used to directly follow the progress for conversion of 3-methoxypyrrole to 3-hydroxypyrrole.

Fig. 10 shows quantitative DNase I footprint titration experiments with the polyamides ImPyPy-γ-PyHpPy-β-Dp and ImHpPy-γ-PyPyPy-β-Dp on the 3'-^{32}P labeled 370-bp pDEH1 *Eco*RI/*Pvu*II restriction fragment. Intact lane, labeled restriction fragment no polyamide or DNase I added; lanes 1-10, DNase I digestion products in the presence of 10 μM, 5 μM, 2 μM, 1 μM, 500 nM, 200 nM, 100 nM, 50 nM, 20 nM, 10 nM ImPyPy-γ-PyPyPy-β-Dp, respectively or 1 μM, 500 nM, 200 nM, 100 nM, 50 nM, 20 nM, 10 nM, 5 nM, 2 nM, 1 nM ImHpPy-γ-PyPyPy-β-Dp, respectively; DNase I lane, DNase I digestion products in the absence of polyamide; A lane, adenine-specific chemical sequencing. Iverson, B. L. & Dervan, P. B. describes an adenine-specific DNA chemical sequencing reaction. *Methods Enzymol.* 15, 7823-7830 (1987). All reactions were done in a total volume of 40 μL. A polyamide stock solution or H_2O was added to an assay buffer containing radiolabeled restriction fragment, with the final solution conditions of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0. Solutions were allowed to equilibrate for 4-12 h at 22 °C before initiation of footprinting reactions. Footprinting reactions, separation of cleavage products, and data analysis were carried out as described. White, S., Baird, E. E. & Dervan, P. describe the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chemistry & Biology* 4, 569-578 (1997).

Figure 11 shows the synthesis of a bifunctional polyamide which incorporates the Hp/Py pair. Treatment of a sample of ImImPyPy-γ-ImHpPyPy-β-Pam-resin (see Figure 6) with 3,3'-diamino-*N*-methyldipropylamine, 55°C, 18 h followed by reverse phase HPLC purification provides the Op polyamide with a free primary amine group which can be coupled to an activated carboxylic acid derivative. Treatment with (i) EDTA-dianhydride, DMSO/NMP, DIEA, 55 °C; (ii) 0.1M NaOH, followed by reverse phase HPLC purification provides the Op-Py-Im-polyamide-EDTA conjugate. Treatment of the 3-methyoxypyrrole polyamide with

thiophenol, NaH, DMF, at 100 °C for 120 min provides polyamide **2** after reverse phase HPLC purification.

Figure 12 shows the determination of the binding orientation of hairpin polyamides ImImPyPy-γ-ImHpPyPy-β-Dp-EDTA•Fe(II) **2-E•Fe(II)** and ImImHpPy-γ-ImPyPyPy-β-Dp-EDTA•Fe(II) **3-E•Fe(II)** by affinity cleaving footprint titration. Top and bottom left: Affinity cleavage experiments on a 3' ^{32}P labeled 250-bp pJK6 *Eco*RI/ *Pvu* II restriction fragment. The 5'-TGGACA-3' and 5'-TGGTCA-3' sites are shown on the right side of the autoradiogram. Top left: lane 1, adenine-specific chemical sequencing reaction; lanes 2-6, 6.5 μM, 1.0 μM, 100 nM, 10 nM, 1 nM polyamide **2-E•Fe(II)**; lane 7, intact restriction fragment, no polyamide added. Bottom left: lane 1, A reaction; lanes 2-6, 8.5 μM, 1.0 μM, 100 nM, 10 nM, 1 nM polyamide **3-E•Fe(II)**; lane 7, intact DNA. All reactions were carried out in a total volume of 40 μL. A stock solution of polyamide or H_20 was added to a solution containing 20 kcpm labeled restriction fragment, affording final solution conditions of 25 mM Tris-Acetate, 20 mM NaCl, 100 μM/ bp calf thymus DNA, at pH 7.0. Solutions were allowed to equilibrate for a minimum of 4 h at 22°K before initiation of reactions. Affinity cleavage reactions were carried out as described White, S., Baird, E.E. & Dervan, P.B. Effects of the A•T/T•A degeneracy of pyrrole-imidazole polyamide recognition in the minor groove of DNA. *Biochemistry* **35**, 6147-6152 (1996). Top and bottom right: Affinity cleavage patterns of **2-E•Fe(II)** and **3-E•Fe(II)** at 100 nM bound to 5'-TGGACA-3' and 5'-TGGTCA-3'. Bar heights are proportional to the relative cleavage intensities at each base pair. Shaded and nonshaded circles denote imidazole and pyrrole carboxamides, respectively. Nonshaded diamonds represent the β-alanine moiety. A curved line represents the γ-aminobutyric acid, and the + represents the positively charged dimethylaminopropylamide tail group. The boxed **Fe** denotes the EDTA•Fe(II) cleavage moiety.

Figure 13 shows quantitative DNase I footprint titration experiments with the polyamides ImPyPyPyPy-γ-ImPyPyPyPyPy-β-Dp and ImHpPyPyPy-γ-ImHpPyPyPy-β-Dp on the 3' ^{32}P labeled 252-bp pJK7 *Eco*RI/ *Pvu* II restriction fragment. For ImPyPyPyPy-γ-ImPyPyPyPy-β-Dp gel (left): lane 1, DNase I digestion products in the absence of polyamide; lanes 2-18, DNase I digestion products in the presence of 1.0 μM, 500 nM, 200, 100, 65, 40, 25, 15, 10, 6.5, 4.0, 2.5, 1.5, 1.0, 0.5, 0.2, 0.1 nM polyamide; lane 19, DNase I digestion products in the absence of polyamide; lane 20, intact restriction fragment; lane 21, guanine-specific chemical sequencing reaction; lane 22, adenine-specific chemical sequencing reaction. For ImHpPyPyPyPy-γ-ImHpPyPyPyPy-β-Dp gel (right): lane 1, intact DNA; lane 2, DNase I digestion products in the absence of polyamide; lanes 3-19, 1.0 μM, 500 nM, 200, 100, 50, 20, 10, 5, 2, 1, 0.5, 0.2. 0.1, 0.05, 0.01, 0.005, 0.001 nM polyamide; lane 20, DNase I digestion products in the absence of polyamide; lane 21, A reaction. All reactions were done in a total volume of 400

μL. A polyamide stock solution or H_20 was added to an assay buffer containing radiolabeled restriction fragment, with the final solution conditions of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0. Solutions were allowed to equilibrate for 4-12 h at 22°C before initiation of footprinting reactions. Footprinting reactions, separation of cleavage products, and data analysis were carried as described. White, S., Baird, E.E. & Dervan, P.B. Effects of the A•T/T•A degeneracy of pyrrole-imidazole polyamide recognition in the minor groove of DNA. *Biochemistry* 35, 6147-6152 (1996).

Fig. 14 shows the 8-ring Hp-Py-Im-polyamide hairpins described by the pairing code of the present invention. The eight ring hairpin template is shown at the top. A polyamide having the formula $X_1X_2X_3X_4$-γ-$X_5X_6X_7X_8$ wherein γ is the -NH-CH_2-CH_2-CH_2-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R–2,4-diaminobutyric acid; X_4/X_5, X_3/X_6, X_2/X_7, and X_1/X_8 represent carboxamide binding pairs which bind the DNA base pairs. The minor groove sequence to be bound is represented as 5'-WGTNNW-3', where the 5'-GTNN-3' core sequence is defined as position a, b, c, and d (W = A or T, N = A, G, C, or T). A linear sequence of aromatic amino acids fills the hairpin template in order to satisfy the ring pairing requirements to correspond to the DNA base pairs in the minor groove to be bound. The ring pairing code as applied is listed in Table 2. The 16 unique hairpin polyamides which target 16 5'-WGTNNW-3' sequences are drawn as binding models where filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole, and the curved line connecting the polyamide subunits represents γ-aminobutyric acid.

Fig. 15 shows the 8-ring Hp-Py-Im-polyamide hairpins described by the pairing code of the present invention. The eight ring hairpin template is shown at the top. A polyamide having the formula $X_1X_2X_3X_4$-γ-$X_5X_6X_7X_8$ wherein γ is the -NH-CH_2-CH_2-CH_2-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R–2,4-diaminobutyric acid; X_4/X_5, X_3/X_6, X_2/X_7, and X_1/X_8 represent carboxamide binding pairs which bind the DNA base pairs. The minor groove sequence to be bound is represented as 5'-WGANNW-3', where the 5'-GANN-3' core sequence is defined as position a, b, c, and d (W = A or T, N = A, G, C, or T). A linear sequence of aromatic amino acids fills the hairpin template in order to satisfy the ring pairing requirements to correspond to the DNA base pairs in the minor groove to be bound. The ring pairing code as applied is listed in Table 2. The 16 unique hairpin polyamides which target 16 5'-WGANNW-3' sequences are drawn as binding models where filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole, and the curved line connecting the polyamide subunits represents γ-aminobutyric acid.

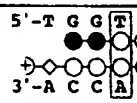
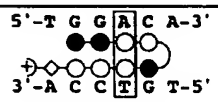
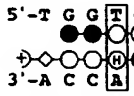
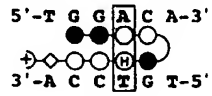
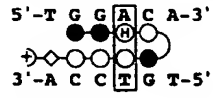
Fig. 16 shows the 8-ring Hp-Py-Im-polyamide hairpins described by the pairing code of the present invention. The eight ring hairpin template is shown at the top. A polyamide having the formula $X_1X_2X_3X_4$-γ-$X_5X_6X_7X_8$ wherein γ is the -NH-CH_2-CH_2-CH_2-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R–2,4-diaminobutyric acid; X_4/X_5, X_3/X_6, X_2/X_7, and X_1/X_8 represent carboxamide binding pairs which bind the DNA base pairs. The minor groove sequence to be bound is represented as 5'-WGGNNW-3', where the 5'-GGNN-3' core sequence is defined as position a, b, c, and d (W = A or T, N = A, G, C, or T). A linear sequence of aromatic amino acids fills the hairpin template in order to satisfy the ring pairing requirements to correspond to the DNA base pairs in the minor groove to be bound. The ring pairing code as applied is listed in Table 2. The 16 unique hairpin polyamides which target 16 5'-WGGNNW-3' sequences are drawn as binding models where filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole, and the curved line connecting the polyamide subunits represents γ-aminobutyric acid.

Fig. 17 shows the 8-ring Hp-Py-Im-polyamide hairpins described by the pairing code of the present invention. The eight ring hairpin template is shown at the top. A polyamide having the formula $X_1X_2X_3X_4$-γ-$X_5X_6X_7X_8$ wherein γ is the -NH-CH_2-CH_2-CH_2-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R–2,4-diaminobutyric acid; X_4/X_5, X_3/X_6, X_2/X_7, and X_1/X_8 represent carboxamide binding pairs which bind the DNA base pairs. The minor groove sequence to be bound is represented as 5'-WGCNNW-3', where the 5'-GCNN-3' core sequence is defined as position a, b, c, and d (W = A or T, N = A, G, C, or T). A linear sequence of aromatic amino acids fills the hairpin template in order to satisfy the ring pairing requirements to correspond to the DNA base pairs in the minor groove to be bound. The ring pairing code as applied is listed in Table 2. The 16 unique hairpin polyamides which target 16 5'-WGCNNW-3' sequences are drawn as binding models where filled and unfilled circles represent imidazole and pyrrole rings respectively; circles containing an H represent 3-hydroxypyrrole, and the curved line connecting the polyamide subunits represents γ-aminobutyric acid.

Four-ring polyamide subunits, covalently coupled to form eight-ring hairpin structures, bind specifically to 6-bp target sequences at subnanomolar concentrations. Trauger, J.W., Baird, E. E. & Dervan, P.B. describe the recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **382**, 559-561 (1996); Swalley, S. E., Baird, E. E. & Dervan, P. B. describe the discrimination of 5'-GGGG-3', 5'-GCGC-3', and 5'-GGCC'3' sequences in the minor groove of DNA by eight-ring hairpin polyamides. *J. Am. Chem. Soc.* **119**, 6953-6961 (1997). The DNA-binding affinities of three eight-ring hairpin polyamides shown in Figure 1 as compound **1**, **2**, and **3** containing pairings of Im/Py, Py/Im opposite G•C,

C•G and either Py/Py, Hp/Py, or Py/Hp at a common single point opposite T•A and A•T has been determined. Equilibrium dissociation constants (K_d) for ImImPyPy-γ-ImPyPyPy-β-Dp **1**, ImImPyPy-γ-ImHpPyPy-β-Dp **2**, ImImHpPy-γ-ImPyPyPy-β-Dp **3** of Figure 1 are shown in Table 1. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. describe a quantitative DNase footprint titration method for studying protein-DNA interactions. *Methods Enzymol.* **130**, 132-181 (1986); The K_d values were determined by quantitative DNase I footprint titration experiments: on a 3' ^{32}P-labeled 250-bp DNA fragment containing the target sites, 5'-TGGACA-3' and 5'-TGGTCA-3' which differ by a single A•T base pair in the fourth position. The DNase footprint gels are shown in Figure 3.

TABLE 1 Equilibrium dissociation constants*

	Polyamide†	5'-TGGTCA-3'	5'-TGGACA-3'	K_{rel}‡
1	Py/Py	 K_d = 0.077 nM	 K_d = 0.15 nM	2.0
2	Py/Hp	 K_d = 15 nM	 K_d = 0.83 nM	0.06
3	Hp/Py	K_d = 0.48 nM	 K_d =37 nM	77

*The reported dissociation constants are the average values obtained from three DNase I footprint titration experiments. The standard deviation for each data set is less than 15% of the reported number. Assays were carried out in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C.
†Ring pairing opposite T•A and A•T in the fourth position.
‡Calculated as K_d(5'-TGGACA-3')/K_d(5'-TGGTC A-3').

Based on the pairing rules for polyamide-DNA complexes both of these sequences are a match for control polyamide **1** which places a Py/Py pairing opposite A•T and T•A at both sites. It was determined that in polyamide **1** (Py/Py) binds to 5'-TGGTCA-3' and 5'-TGGACA-3' within a factor of 2 (K_d = 0.077 or 0.15 nM respectively). In contrast, polyamide **2** (Py/Hp) binds to 5'-TGGTCA-3' and 5'-TGGACA-3' with dissociation constants which differ by a factor of 18 (K_d = 15 nM and 0.83 nM respectively). By reversing the pairing in polyamide **3** (Hp/Py) the dissociation constants differ again in the opposite direction by a factor of 77 (K_D = 0.48 nM and 37 nM respectively. Control experiments performed on separate DNA fragments; reveal that neither a 5'-TGGGCA-3' or a 5'-TGGCCA-3' site is bound by polyamide **2** or **3** at concentrations ≤ 100 nM, indicating that the Hp/Py and Py/Hp ring pairings do not bind opposite G•C or C•G. The A•T vs. T•A discrimination is achieved preferably when the two neighboring base pairs are G•C and C•G (GTC vs. GAC).

The specificity of polyamides 2 and 3 for sites which differ by a single A•T/T•A base pair results from small chemical changes. Replacing the Py/Py pair in 1 with a Py/Hp pairing as in 2, a single substitution of C3-OH for C3-H, destabilizes interaction with 5'-TGGTCA-3' by 191-fold, a free energy difference of 3.1 kcal mol^{-1}. Interaction of 2 with 5'-TGGACA-3' is destabilized only 6-fold relative to 1, a free energy difference of 1.1 kcal mol^{-1}. Similarly, replacing the Py/Py pair in 1 with Hp/Py as in 3 destabilizes interaction with 5'-TGGACA-3' by 252-fold, a free energy difference of 3.2 kcal mol^{-1}. Interaction of 3 with 5'TGGTCA-3' is destabilized only 6-fold relative to 1, a free energy difference of 1.0 kcal mol^{-1}.

The polyamides of this invention provide for coded targeting of predetermined DNA sequences with affinity and specificity comparable to sequence-specific DNA binding proteins. Hp, Im, and Py polyamides complete the minor groove recognition code using three aromatic amino acids which combine to form four ring pairings (Im/Py, Py/Im, Hp/Py, and Py/Hp) which complement the four Watson-Crick base pairs, as shown in TABLE 2. There are a possible 240 four base pair sequences which contain at least 1 A•T or T•A base pair and therefore can advantageously use an Hp/Py, or Py/Hp carboxamide binding. Polyamides binding to any of these sequences can be designed in accordance with the code of TABLE 2.

TABLE 2 Pairing code for minor groove recognition*

Pair	G•C	C•G	T•A	A•T
Im/Py	+	-	-	-
Py/Im	-	+	-	-
Hp/Py	-	-	+	-
Py/Hp	-	-	-	+

* favored (+), disfavored (-)

For certain G•C rich sequences the affinity of polyamide•DNA complexes may be enhanced by substitution of an Im/β pair for Im/Py at G•C and β/Im for Py/Im at C•G. At A•T and T•A base pairs, either a Py/β, β/Py, and β/β may be used. The alternate aliphatic/aromatic amino acid pairing code is described in Table 3.

TABLE 3 Aliphatic/Aromatic substitution for ring pairings*

Pair	Substitution
Im/Py	Im/β
Py/Im	β/Im
Hp/Py	Py/β, β/Py, Hp/β, β/β
Py/Hp	Py/β, β/Py, β/Hp, β/β

U. S. Patent 5,578,444 describes numerous promoter region targeting sequences from which base pair sequences for targeting a polyamide can be identified.

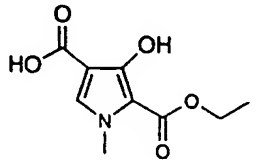
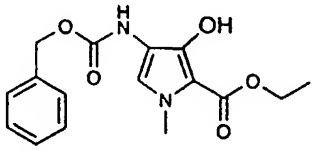
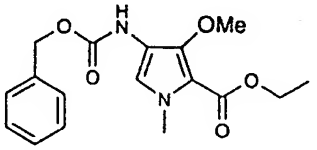
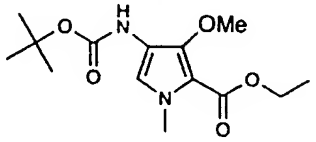
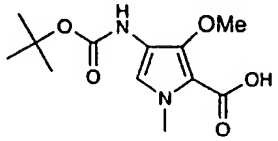
PCT U.S. 97/003332 describes methods for synthesis of polyamides which are suitable for preparing polyamides of this invention. The use of β-alanine in place of a pyrrole amino acid in the synthetic methods provides aromatic/aliphatic pairing (Im/β, β/Im, Py/β, and β/Py) and aliphatic/aliphatic pairing (β/β) substitution. The use of γ-aminobutyric acid, or a substituted γ-aminobutyric acid such as (R)-2,4 diaminobutyric acid, provides for preferred hairpin turns. The following examples illustrate the synthesis of polyamides of the present invention.

Example 1:

PREPARATION OF A PROTECTED Hp MONOMER FOR SOLID PHASE SYNTHESIS.

Distamycin and its analogs have previously been considered targets of traditional multistep synthetic chemistry. Arcamone, F., Orezzi, P. G., Barbieri, W., Nicolella, V. & Penco, S. describe a solution phase synthesis of distamycin *Gazz. Chim. Ital.* **1967**, *97*, 1097. The repeating amide of distamycin is formed from an aromatic carboxylic acid and an aromatic amine. The aromatic acid is often unstable to decarboxylation, and the aromatic amines have been found to be air and light sensitive. Lown, J. W. & Krowicki, K. describe a solution phase synthesis of Distamycin *J. Org. Chem.* **1985**, *50*, 3774. The variable coupling yields, long reaction times (often >24 h), numerous side products, and reactive intermediates (acid chlorides and trichloro ketones) characteristic of the traditional solution phase coupling reactions make the synthesis of the aromatic carboxamides problematic. B. Merrifield describes the solid phase synthesis of a tetrapeptide *J. Am. Chem. Soc.* **1963**, 85, 2149. In order to implement an efficient solid phase methodology for the synthesis of the pyrrole- imidazole polyamides, the following components were developed: (1) a synthesis which provides large quantities of appropriately protected monomer or dimer building blocks in high purity, (2) optimized protocols for forming an amide in high yield from a support-bound aromatic amine and an aromatic carboxylic acid, (3) methods for monitoring reactions on the solid support, and (4) a stable resin linkage agent that can be cleaved in high yield upon completion of the synthesis. Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. In order to prepare polyamides which contain the 3-hydroxypyrrole monomer, a synthesis has been developed which allows the appropriately protected Boc-Op acid monomer to be prepared on 50 g scale. ^{1}H NMR and ^{13}C NMR spectra were recorded on a General Electric-QE 300 NMR spectrometer in CD_3OD or DMSO-d_6, with chemical shifts reported in parts per million relative to residual CHD_2OD or DMSO-d_5, respectively. IR spectra were recorded on a Perkin-Elmer FTIR spectrometer. High-resolution mass spectra were recorded using fast atom bombardment (FABMS) techniques at the Mass Spectrometry Laboratory at the University of California, Riverside. Reactions were executed under an inert argon atmosphere. Reagent grade chemicals were used as received unless otherwise noted. Still, W. C., Kahn, M. & Mitra, A. describe flash column chromatography *J. Org. Chem.* **1978**, *40*, 2923-2925. Flash chromatography was carried out using EM science Kieselgel 60 (230-400) mesh. Thin-layer chromatography was performed on EM Reagents silica gel plates (0.5 mm thickness). All compounds were visualized with short-wave ultraviolet light.

Table 4 :Intermediates for preparation of Boc-protected 3-methoxypyrrole

NAME	STRUCTURE
Ethyl 4-carboxy-3-hydroxy-1-methylpyrrole-2-carboxylate.	
Ethyl 4-[(Benzyloxycarbonyl)amino]-3-hydroxy-1-methylpyrrole-2-carboxylate	
Ethyl 4-[(Benzyloxycarbonyl)amino]-3-methoxy-1-methylpyrrole-2-carboxylate	
Ethyl 4-[(tert-Butyloxycarbonyl)amino]-3-methoxy-1-methylpyrrole-2-carboxylate	
4-[(tert-Butyloxycarbonyl)amino]-3-methoxy -1-methylpyrrole-2-carboxylic acid	

Ethyl 4-[(benzyloxycarbonyl)amino]-3-hydroxy-1-methylpyrrole-2-carboxylate Ethyl-4-carboxy-3-hydroxy-1-methylpyrrole-2-carboxylate (60 g, 281.7 mmol) was dissolved in 282 mL acetonitrile. TEA (28.53 g, 282 mmol) was added, followed by diphenylphosphorylazide (77.61 g, 282 mmol). The mixture was refluxed for 5 hours, followed by addition of benzyl alcohol (270 ml) and reflux continued overnight. The solution was cooled and volitiles removed *in vacuo*. The residue was absorbed onto silca and chromatagraphed, 4:1 hexanes : ethyl acetate, to give a white solid (21.58 g, 24%) ^{1}H NMR (DMSO-d$_6$) δ 8.73 (s, 1H), 8.31 (s, 1H), 7.31 (m, 5H), 6.96 (s, 1H), 5.08 (s, 2H), 4.21 (q, 2H, J = 7.1 Hz), 3.66 (s, 3H), 1.25 (t, 3H, J = 7.1 Hz); MS *m/e* 319.163 (M+H 319.122 calcd. for $C_{16}H_{18}N_2O_5$).

Ethyl 4-[(tert-butoxycarbonyl)amino]-3-methoxy-1-methylpyrrole-2- carboxylate. Ethyl 4-[(benzyloxycarbonyl)amino]-3-hydroxy-1-methylpyrrole-2-carboxylate (13.4 g, 42.3 mmol) was dissolved in 110 mL acetone. Anhydrous K_2CO_3 (11.67 g, 84.5 mmol) was added,

followed by methyliodide (5.96 g, 42.3 mmol) and dimethylaminopyridine (0.5 g, 4.23 mmol) and the mixture stirred overnight. The solid K_2CO_3 was removed by filtration and 200 ml water added. Volitiles were removed *in vacuo* and the solution made acidic with addition of 1*N* H_2SO_4 . The aqueous layer was extracted with diethyl ether. Organic layers were combined, washed with 10% H_2SO_4, dried over $MgSO_4$, and dried to give a white solid. The solid was used without further purification and dissolved in 38 ml DMF. DIEA (11 ml), Boc anhydride (9.23 g, 42.3 mmol), and 10 % Pd/C (500 mg) were added and the solution stirred under hydrogen (1 atm) for 2.1 h. The slurry was filtered through celite which was washed with methanol. Water 250 ml was added and volitiles removed *in vacuo*. The aqueous layer was extracted with ether. Organic layers were combined, washed with water and brine, and dried over $MgSO_4$. Solvent was removed *in vacuo* to give a white solid (8.94 g, 71%) ^{1}H NMR (DMSO-d_6) δ 8.43 (s, 1H), 7.03 (s, 1H), 4.19 (q, 2H, J = 7.1 Hz), 3.70 (s, 3H), 3.67 (s, 3H), 1.42 (s, 9H), 1.26 (t, 3H, J = 7.1); MS *m/e* 299.161 (M+H 299.153 calcd. for $C_{14}H_{22}N_2O_5$).

Ethyl 4-[(benzyloxycarbonyl)amino]-3-hydroxy-1-methylpyrrole-2-carboxylate Ethyl 4-[(*tert*-butoxycarbonyl)amino]-3-methoxy-1-methylpyrrole-2-carboxylate (9.0 g, 30.2 mmol) was dissolved in 30 mL ethanol. NaOH (30 ml, 1 M, aq) was added and the solution stirred for 4 days. Water (200 ml) was added and ethanol removed *in vacuo*. The solution was extracted with diethyl ether, aqueous layer acidified to pH = 2-3, and extracted again with diethyl ether. Organic layers were dried over $MgSO_4$, and solvent removed *in vacuo* to give a white solid (6.0 g, 20.5 mmol, 87% based on recovered SM) ^{1}H NMR (DMSO-d_6) δ 12.14 (s, 1H), 8.37 (s, 1H), 6.98 (s, 1H), 3.69 (s, 3H), 3.66 (s, 3H), 1.42 (s, 9H); MS *m/e* 293.112 (M+H 293.104 calcd. for $C_{12}H_{18}N_2O_5$).

EXAMPLE 2:
SOLID PHASE SYNTHESIS OF 3-HYDROXYPYRROLE POLYAMIDES.

Cycling protocols were optimized to afford high stepwise coupling yields (>99%). Deprotection by aminolysis affords up to 100 mg quantities of polyamide. Solid phase polyamide synthesis protocols were modified from the *in situ* neutralization Boc-chemistry protocols. Schnolzer, M., Alewood, P., Jones, A., Alewood, D., Kent, S.B.H. report rapid *in situ* neutralization for solid phase peptide synthesis *Int. J. Peptide. Protein. Res.* **1992**, *40*, 180. Coupling cycles are rapid, 72 min per residue for manual synthesis or 180 min per residue for machine-assisted synthesis, and require no special precautions beyond those used for ordinary solid phase peptide synthesis. Manual solid phase synthesis of a pyrrole-imidazole polyamide consists of a dichloromethane (DCM) wash, removal of the Boc group with trifluoroacetic acid (TFA)/DCM/thiophenol (PhSH), a DCM wash, a DMF wash, taking a resin sample for analysis, addition of activated monomer, addition of DIEA if necessary, coupling for 45 min, taking a

resin sample for analysis, and a final DMF wash (Figure 5, Table I). In addition, the manual solid phase protocol for synthesis of pyrrole-imidazole polyamides has been adapted for use on a ABI 430A peptide synthesizer. The aromatic amine of the pyrrole and imidazole do not react in the quantitative ninhydrin test. Stepwise cleavage of a sample of resin and analysis by HPLC indicates that high stepwise yields (> 99%) are routinely achieved.

Dicyclohexylcarbodiimide (DCC), Hydroxybenzotriazole (HOBt), 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexa-fluorophosphate (HBTU) and 0.2 mmol/gram Boc-β-alanine-(-4-carboxamidomethyl)-benzyl-ester-copoly(styrene-divinylbenzene) resin (Boc-β-Pam-Resin) was purchased from Peptides International (0.2 mmol/gram), NovaBiochem (0.6 mmol/gram), or Peninsula (0.6 mmol/gram). ((*R*)-2-Fmoc-4-Boc-diaminobutyric acid, (*S*)-2-Fmoc-4-Boc-diaminobutyric acid, and (*R*)-2-amino-4-Boc-diaminobutyric acid were purchased from Bachem. *N,N*-diisopropylethylamine (DIEA), *N,N*-dimethylformamide (DMF), *N*-methylpyrrolidone (NMP), DMSO/NMP, Acetic anhydride (Ac_2O), and 0.0002 M potassium cyanide/pyridine were purchased from Applied Biosystems. Dichloromethane (DCM) and triethylamine (TEA) were reagent grade from EM, thiophenol (PhSH), dimethylaminopropylamine (Dp), Sodium Hydride, (*R*)-α-methoxy-α-(trifuoromethyl)phenylacetic acid ((*R*)MPTA) and (*S*)-α-methoxy-α-(trifouromethyl)phenylacetic acid ((*S*)MPTA) were from Aldrich, trifluoroacetic acid (TFA) Biograde from Halocarbon, phenol from Fisher, and ninhydrin from Pierce. All reagents were used without further purification.

Quik-Sep polypropylene disposable filters were purchased from Isolab Inc. ^{1}H NMR spectra were recorded on a General Electric-QE NMR spectrometer at 300 MHz with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrophotometer. Optical rotations were recorded on a JASCO Dip 1000 Digital Polarimeter. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed at the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed on either a HP 1090M analytical HPLC or a Beckman Gold system using a RAINEN C_{18}, Microsorb MV, 5μm, 300 x 4.6 mm reversed phase column in 0.1% (wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory reverse phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C18 column equipped with a guard, 0.1% (wt/v) TFA, 0.25% acetonitrile/min. 18MΩ water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered.

Activation of Boc-3-methoxypyrrole acid. The amino acid (0.5 mmol) was dissolved in 2 mL DMF. HBTU (190 mg, 0.5 mmol) was added followed by DIEA (1 mL) and the resulting mixture was shaken for 5 min.

Activation of Imidazole-2-carboxylic acid, γ-aminobutyric acid, Boc-glycine, and Boc-β-alanine. The appropriate amino acid or acid (2 mmol) was dissolved in 2 mL DMF. HBTU (720 mg, 1.9 mmol) was added followed by DIEA (1 mL) and the solution shaken for at least 5 min.

Activation of Boc-Imidazole acid. Boc imidazole acid (257 mg, 1 mmol) and HOBt (135 mg, 1 mmol) were dissolved in 2 mL DMF, DCC (202 mg, 1 mmol) is then added and the solution allowed to stand for at least 5 min.

Acetylation Mix. 2 mL DMF, DIEA (710 μL, 4.0 mmol), and acetic anhydride (380 μL, 4.0 mmol) were combined immediately before use.

Manual Synthesis Protocol. Boc-ß-alanine-Pam-Resin (1.25 g, 0.25 mmol) is placed in a 20 mL glass reaction vessel, shaken in DMF for 5 min and the reaction vessel drained. The resin was washed with DCM (2 x 30 s.) and the Boc group removed with 80% TFA/DCM/0.5M PhSH, 1 x 30s., 1 x 20 min The resin was washed with DCM (2 x 30 s.) followed by DMF (1 x 30 s.) A resin sample (5 - 10 mg) was taken for analysis. The vessel was drained completely and activated monomer added, followed by DIEA if necessary. The reaction vessel was shaken vigorously to make a slurry. The coupling was allowed to proceed for 90 min, and a resin sample taken. Acetic anhydride (1 mL) was added and the reaction shaken for 5 min. The reaction vessel was then washed with DMF, followed by DCM.

Machine-Assisted Protocols. Machine-assisted synthesis was performed on a ABI 430A synthesizer on a 0.18 mmol scale (900 mg resin; 0.2 mmol/gram). Each cycle of amino acid addition involved: deprotection with approximately 80% TFA/DCM/0.4M PhSH for 3 minutes, draining the reaction vessel, and then deprotection for 17 minutes; 2 dichloromethane flow washes; an NMP flow wash; draining the reaction vessel; coupling for 1 hour with *in situ* neutralization, addition of dimethyl sulfoxide (DMSO)/NMP, coupling for 30 minutes, addition of DIEA, coupling for 30 minutes; draining the reaction vessel; washing with DCM, taking a resin sample for evaluation of the progress of the synthesis by HPLC analysis; capping with acetic anhydride/DIEA in DCM for 6 minutes; and washing with DCM. A double couple cycle is employed when coupling aliphatic amino acids to imidazole, all other couplings are performed with single couple cycles.

The ABI 430A synthesizer was left in the standard hardware configuration for NMP-HOBt protocols. Reagent positions 1 and 7 were DIEA, reagent position 2 was TFA/0.5M thiophenol, reagent position 3 was 70% ethanolamine/methanol, reagent position 4 was acetic anhydride, reagent position 5 was DMSO/NMP, reagent position 6 was methanol, and reagent position 8 was DMF. New activator functions were written, one for direct transfer of the cartridge contents to the concentrator (switch list 21, 25, 26, 35, 37, 44), and a second for transfer of reagent position 8 directly to the cartridge (switch list 37, 39, 45, 46).

Boc-Py-OBt ester (357 mg, 1 mmol) was dissolved in 2 ml DMF and filtered into a synthesis cartridge. Boc-Im acid monomer was activated (DCC/HOBt), filtered, and placed in a synthesis cartridge. Imidazole-2-carboxylic acid was added manually. At the initiation of the coupling cycle the synthesis was interrupted, the reaction vessel vented and the activated monomer added directly to the reaction vessel through the resin sampling loop via syringe. When manual addition was necessary an empty synthesis cartridge was used. Aliphatic amino acids (2 mmol) and HBTU (1.9 mmol) were placed in a synthesis cartridge. 3 ml of DMF was added using a calibrated delivery loop from reagent bottle 8, followed by calibrated delivery of 1 ml DIEA from reagent bottle 7, and a 3 minute mixing of the cartridge.

The activator cycle was written to transfer activated monomer directly from the cartridge to the concentrator vessel, bypassing the activator vessel. After transfer, 1 ml of DIEA was measured into the cartridge using a calibrated delivery loop, and the DIEA solution combined with the activated monomer solution in the concentrator vessel. The activated ester in 2:1 DMF/DIEA was then transferred to the reaction vessel. All lines were emptied with argon before and after solution transfers.

ImImOpPy-γ-ImPyPyPy-β-Dp. ImImOpPy-γ-ImPyPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.66 mmol/g). Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. 3-hydroxypyrrole-Boc-amino acid (0.7 mmol) was incorporated by placing the amino acid (0.5 mmol) and HBTU (0.5 mmol) in a machine synthesis cartridge. Upon automated delivery of DMF (2 mL) and DIEA (1 mL) activation occurs. A sample of ImImOpPy-γ-ImPyPyPy-β-Pam-Resin (400 mg, 0.40 mmol/gram) was placed in a glass 20 mL peptide synthesis vessel and treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImImOpPy-γ-ImPyPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (97 mg, 49% recovery). UV (H_2O) λ_{max} 246, 316 (66,000); ^{1}H NMR (DMSO-d_6)

δ 10.24 (s, 1 H), 10.14 (s, 1 H), 9.99 (s, 1 H), 9.94 (s, 1 H), 9.88 (s, 1 H), 9.4 (br s, 1 H), 9.25 (s, 1 H), 9.11 (s, 1 H), 8.05 (m, 3 H), 7.60 (s, 1 H), 7.46 (s, 1 H), 7.41 (s, 1 H), 7.23 (d, 1), 7.21 (d, 1 H), 7.19 (d, 1 H), 7.13 (m, 2 H), 7.11 (m, 2 H), 7.02 (d, 1 H), 6.83 (m, 2 H), 3.96 (s, 6 H), 3.90 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.75 (d, 9 H), 3.33 (q, 2 H, J = 5.4 Hz), 3.15 (q, 2 H, J = 5.5 Hz), 3.08 (q, 2 H, J = 6.0 Hz), 2.96 (quintet, 2 H, J = 5.6 Hz), 2.70 (d, 6 H, J = 4.5 Hz), 2.32 (m, 4 H), 1.71 (m, 4 H); MALDI-TOF-MS (monoisotopic), 1253.5 (1253.6 calc. for $C_{58}H_{72}N_{22}O_{11}$).

ImImHpPy-γ-ImPyPyPy. In order to remove the methoxy protecting group, a sample of ImImOpPy-γ-ImPyPyPy-β-Dp (5 mg, 3.9 μmol) was treated with sodium thiophenoxide at 100 °C for 2 h. DMF (1000 μL) and thiophenol (500 μL) were placed in a (13 x 100 mm) disposable Pyrex screw cap culture tube. A 60 % dispersion of sodium hydride in mineral oil (100 mg) was *slowly* added. Upon completion of the addition of the sodium hydride, ImImOpPy-γ-ImPyPyPy-β-Dp (5 mg) dissolved in DMF (500 μL) was added. The solution was agitated, and placed in a 100 °C heat block, and deprotected for 2 h. Upon completion of the reaction the culture tube was cooled to 0°C, and 7 ml of a 20 % (wt/v) solution of trifluoroacetic acid added. The aqueous layer is separated from the resulting biphasic solution and purified by reversed phase HPLC. ImImHpPy-γ-ImPyPyPy-β-Dp is recovered as a white powder upon lyophilization of the appropriate fractions (3.8 mg, 77 % recovery). UV (H_2O) λ_{max} 246, 312 (66,000); ^{1}H NMR (DMSO-d_6) δ 10.34 (s, 1 H), 10.24 (s, 1 H), 10.00 (s, 2 H), 9.93 (s, 1 H), 9.87 (s, 1 H), 9.83 (s, 1 H), 9.4 (br s, 1 H), 9.04 (s, 1 H), 8.03 (m, 3 H), 7.58 (s, 1 H), 7.44 (s, 1 H), 7.42 (s, 1 H), 7.23 (s, 1 H), 7.20 (m, 3 H), 7.12 (m, 2 H), 7.05 (d, 1 H), 7.02 (d, 1 H), 6.83 (s, 1 H), 6.79 (s, 1 H), 3.96 (s, 6 H), 3.90 (s, 3 H), 3.81 (s, 6 H), 3.79 (s, 3 H), 3.75 (d, 6 H), 3.33 (q, 2 H, J = 5.4 Hz), 3.14 (q, 2 H, J = 5.4 Hz), 3.08 (q, 2 H, J = 6.1 Hz), 2.99 (quintet, 2 H, J = 5.4 Hz), 2.69 (d, 6 H, J = 4.2 Hz), 2.31 (m, 4 H), 1.72 (m, 4 H); MALDI-TOF-MS (monoisotopic), 1239.6 (1239.6 calc. for $C_{57}H_{71}N_{22}O_{11}$).

ImImPyPy-γ-ImOpPyPy-β-Dp. ImImPyPy-γ-ImOpPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.66 mmol/g) as described for ImImOpPy-γ-ImPyPyPy-β-Dp. A sample of ImImPyPy-γ-ImOpPyPy-β-Pam-Resin (400 mg, 0.40 mmol/gram) was placed in a glass 20 mL peptide synthesis vessel and treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImImPyPy-γ-ImOpPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (101 mg, 50% recovery). UV (H_2O) λ_{max} 246, 316 (66,000); MALDI-TOF-MS (monoisotopic), 1253.6 (1253.6 calc. for $C_{58}H_{72}N_{22}O_{11}$).

ImImPyPy-γ-ImHpPyPy. A sample of ImImPyPy-γ-ImOpPyPy-β-Dp (5 mg, 3.9 μmol) was treated with sodium thiophenoxide and purified by reversed phase HPLC as described for ImImHpPy-γ-ImPyPyPy-β-Dp. ImImPyPy-γ-ImHpPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (3.2 mg, 66 % recovery). UV (H_2O) λ_{max} 246, 312 (66,000); MALDI-TOF-MS (monoisotopic), 1239.6 (1239.6 calc. for $C_{57}H_{71}N_{22}O_{11}$).

ImPyPy-γ-OpPyPy-β-Dp. ImPyPy-γ-OpPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.66 mmol/g). Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. 3-hydroxypyrrole-Boc-amino acid (0.7 mmol) was incorporated by placing the amino acid (0.5 mmol) and HBTU (0.5 mmol) in a machine synthesis cartridge. Upon automated delivery of DMF (2 mL) and DIEA (1 mL) activation occurs. A sample of ImPyPy-γ-OpPyPy-β-Pam-Resin (400 mg, 0.45 mmol/gram) was placed in a glass 20 mL peptide synthesis vessel and treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImPyPy-γ-OpPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (45 mg, 25% recovery). UV (H_2O) λ_{max} 246, 310 (50,000); ^{1}H NMR (DMSO-d_6) δ 10.45 (s, 1 H), 9.90 (s, 1 H), 9.82 (s, 1 H), 9.5 (br s, 1 H), 9.38 (s, 1 H), 9.04 (s, 1 H), 8.02 (m, 3 H), 7.37 (s, 1 H), 7.25 (m, 2 H), 7.15 (d, 1 H, J = 1.6 Hz), 7.11 (m, 2 H), 7.09 (d, 1 H), 7.03 (d, 1 H), 6.99 (d, 1 H), 6.87 (d, 1 H), 6.84 (d, 1 H), 3.96 (s, 3 H), 3.81 (s, 6 H), 3.77 (s, 6 H), 3.76 (s, 3 H), 3.74 (s, 1 H), 3.34 (q, 2 H, J = 5.6 Hz), 3.20 (q, 2 H, J = 5.8 Hz), 3.09 (q, 2 H, J = 6.1 Hz), 2.97 (quintet, 2 H, J = 5.3 Hz), 2.70 (d, 6 H, J = 3.9 Hz), 2.34 (m, 4 H), 1.73 (m, 4 H); MALDI-TOF-MS (monoisotopic), 1007.6 (1007.5 calc. for $C_{48}H_{63}N_{16}O_9$).

ImPyPy-γ-HpPyPy. In order to remove the methoxy protecting group, a sample of ImPyPy-γ-OpPyPy-β-Dp (5 mg, 4.8 μmol) was treated with sodium thiophenoxide at 100 °C for 2 h. DMF (1000 μL) and thiophenol (500 μL) were placed in a (13 x 100 mm) disposable Pyrex screw cap culture tube. A 60 % dispersion of sodium hydride in mineral oil (100 mg) was *slowly* added. Upon completion of the addition of the sodium hydride, ImImPyPy-γ-ImOpPyPy-β-Dp (5 mg) dissolved in DMF (500 μL) was added. The solution was agitated, and placed in a 100 °C heat block, and deprotected for 2 h. Upon completion of the reaction the culture tube was cooled to 0°C, and 7 ml of a 20 % (wt/v) solution of trifluoroacetic acid added. The aqueous layer is separated from the resulting biphasic solution and purified by reversed phase HPLC. ImImHpPy-γ-ImHpPyPy-β-Dp is recovered as a white powder upon lyophilization of the appropriate fractions (2.5 mg, 52 % recovery). UV (H_2O) λ_{max} 246, 310 (50,000); ^{1}H NMR (DMSO-d_6) δ 10.44 (s, 1 H), 10.16 (s, 1 H), 9.90 (s, 1 H), 9.77 (s, 1 H), 9.5 (br s, 1 H), 9.00 (s,

1 H), 8.03 (m, 3 H), 7.37 (s, 1 H), 7.26 (m, 2 H), 7.14 (d, 1 H, J = 1.7 Hz), 7.12 (m, 2 H), 7.02 (d, 1 H), 6.93 (d, 1 H), 6.88 (d, 1 H), 6.82 (d, 1 H), 6.72 (d, 1 H), 3.96 (s, 3 H), 3.81 (s, 6 H), 3.77 (s, 3 H), 3.76 (s, 3 H), 3.74 (s, 1 H), 3.36 (q, 2 H, J = 5.4 Hz), 3.22 (q, 2 H, J = 5.9 Hz), 3.09 (q, 2 H, J = 5.5 Hz), 2.98 (quintet, 2 H, J = 5.3 Hz), 2.70 (d, 6 H, J = 4.3 Hz), 2.34 (m, 4 H), 1.78 (m, 4 H); MALDI-TOF-MS (monoisotopic), 994.2 (993.5 calc. for $C_{47}H_{61}N_{16}O_9$).

Table 5. Mass spectral characterization of Op and Hp polyamides, synthesized and purified as described for ImImOpPy-γ-ImPyPyPy-β-Dp and ImImHpPy-γ-ImPyPyPy-β-Dp.

POLYAMIDE	FORMULA	(M+H)CALCD	FOUND
ImOpPy-γ-PyPyPy-β-Dp	$C_{48}H_{63}N_{16}O_9$	1007.5	1007.5
ImHpPy-γ-PyPyPy-β-Dp	$C_{47}H_{61}N_{16}O_9$	993.5	993.2
ImPyOp-γ-PyPyPy-β-Dp	$C_{48}H_{63}N_{16}O_9$	1007.5	1007.5
ImPyHp-γ-PyPyPy-β-Dp	$C_{47}H_{61}N_{16}O_9$	993.5	993.4
ImPyPy-γ-OpPyPy-β-Dp	$C_{48}H_{63}N_{16}O_9$	1007.5	1007.6
ImPyPy-γ-HpPyPy-β-Dp	$C_{47}H_{61}N_{16}O_9$	993.5	993.2
ImPyPy-γ-PyOpPy-β-Dp	$C_{48}H_{63}N_{16}O_9$	1007.5	1007.5
ImPyPy-γ-PyHpPy-β-Dp	$C_{47}H_{61}N_{16}O_9$	993.5	993.4
ImOpOp-γ-PyPyPy-β-Dp	$C_{49}H_{65}N_{16}O_{10}$	1037.5	1037.5
ImHpHp-γ-PyPyPy-β-Dp	$C_{47}H_{61}N_{16}O_{10}$	1009.5	1009.4
ImImOpPy-γ-ImPyPyPy-β-Dp	$C_{58}H_{72}N_{22}O_{11}$	1253.6	1253.5
ImImHpPy-γ-ImPyPyPy-β-Dp	$C_{57}H_{71}N_{22}O_{11}$	1239.6	1239.6
ImImPyPy-γ-ImOpPyPy-β-Dp	$C_{58}H_{72}N_{22}O_{11}$	1253.6	1253.6
ImImPyPy-γ-ImHpPyPy-β-Dp	$C_{57}H_{71}N_{22}O_{11}$	1239.6	1239.6
ImOpPyPy-γ-ImOpPyPy-β-Dp	$C_{60}H_{76}N_{21}O_{12}$	1282.6	1282.6
ImHpPyPy-γ-ImHpPyPy-β-Dp	$C_{58}H_{72}N_{21}O_{12}$	1254.6	1254.6
ImImOpPy-γ-ImOpPyPy-β-Dp	$C_{59}H_{75}N_{22}O_{12}$	1283.6	1283.6
ImImHpPy-γ-ImHpPyPy-β-Dp	$C_{57}H_{71}N_{22}O_{12}$	1255.6	1255.5
ImOpPyPy-γ-PyPyPyPy-β-Dp	$C_{60}H_{75}N_{20}O_{11}$	1251.6	1251.5
ImPyPyPy-γ-PyPyOpPy-β-Dp	$C_{60}H_{75}N_{20}O_{11}$	1251.6	1251.5
ImImPyPy-γ-ImPyOpPy-β-Dp	$C_{58}H_{72}N_{22}O_{11}$	1253.6	1253.7
ImOpPyPyPy-γ-ImOpPyPyPy-β-Dp	$C_{72}H_{88}N_{25}O_{14}$	1526.7	1526.6
ImHpPyPyPy-γ-ImHpPyPyPy-β-Dp	$C_{70}H_{84}N_{25}O_{14}$	1498.7	1498.0
ImImPyPyPy-γ-ImOpOpPyPy-β-Dp	$C_{71}H_{87}N_{26}O_{14}$	1527.7	1527.7

EXAMPLE 3:
DETERMINATION OF POLYAMIDE BINDING AFFINITY AND SEQUENCE SPECIFICITY.

Representative-footprint titration experiments are shown in Figures 3 and 10. A 252-bp DNA fragment which is typically used for the footprint titration experiments provides 247 possible 6-bp binding sites for an eight-ring hairpin polyamide. Thus, in addition to providing DNA binding affinities, the footprint titration experiments also reveal DNA binding sequence-specificity. The DNA binding sequence specificity of polyamides which differ by a single Py/Py, Hp/Py, or Py/Hp pair for sites which differ by a single A•T or T•A base pair are described in Tables 1, 6, and 7.

Quantitative DNase I Footprint Titrations All reactions were executed in a total volume of 400 μL (Brenowitz, M. *et al.*, 1986). A polyamide stock solution or H_2O (for reference lanes) was added to an assay buffer containing 3'-^{32}P radiolabeled restriction fragment (20,000 cpm), affording final solution conditions of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, and either (i) a suitable concentration range of polyamide, or (ii) no polyamide (for reference lanes). The solutions were allowed to equilibrate for 24 hours at 22°C. Footprinting reactions were initiated by the addition of 10 μL of a stock solution of DNase I (at the appropriate concentration to give ~55% intact DNA) containing 1 mM dithiothreitol and allowed to proceed for 7 minutes at 22°C. The reactions were stopped by the addition of 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 23 μM base pair calf thymus DNA, and 0.6 mg/ml glycogen, and ethanol precipitated. The reactions were resuspended in 1 × TBE/ 80% formamide loading buffer, denatured by heating at 85°C for 15 minutes, and placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% crosslinking, 7 M urea) in 1 × TBE at 2000 V for 1.5 h. Gels were dried on a slab dryer and then exposed to a storage phosphor screen at 22°C.

Photostimuable storage phosphor imaging plates (Kodak Storage Phosphor Screen SO230 obtained from Molecular Dynamics) were pressed flat against dried gel samples and exposed in the dark at 22°C for 12-24 hours. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens (Johnston et al., 1990). The data were analyzed by performing volume integration of the target site and reference blocks using the ImageQuant v. 3.3 software running on a Compaq Pentium 80.

Quantitative DNase I Footprint Titration Data Analysis was performed by taking a background-corrected volume integration of rectangles encompassing the footprint sites and a reference site at which DNase I reactivity was invariant across the titration generated values for

the site intensities (I_{site}) and the reference intensity (I_{ref}). The apparent fractional occupancy (θ_{app}) of the sites were calculated using the equation:

$$\theta_{app} = 1 - \frac{I_{site}/I_{ref}}{I_{site}^{\circ}/I_{ref}^{\circ}} \quad (1)$$

where I_{site}° and I_{ref}° are the site and reference intensities, respectively, from a DNase I control lane to which no polyamide was added.

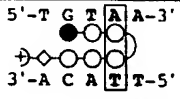
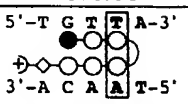
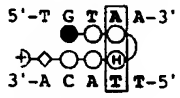
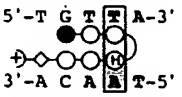
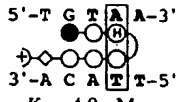
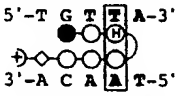
The ($[L]_{tot}$, θ_{app}) data were fit to a Langmuir binding isotherm (eq. 2, n=1) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation:

$$\theta_{fit} = \theta_{min} + (\theta_{max} - \theta_{min}) \frac{K_a^{n}[L]^{n}_{tot}}{1 + K_a^{n}[L]^{n}_{tot}} \quad (2)$$

where $[L_{tot}]$ is the total polyamide concentration, K_a is the equilibrium association constant, and θ_{min} and θ_{max} are the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. The data were fit using a nonlinear least-squares fitting procedure of KaleidaGraph software (v. 3.0.1, Abelbeck Software) with K_a, θ_{max}, and θ_{min} as the adjustable parameters. The goodness of fit of the binding curve to the data points is evaluated by the correlation coefficient, with $R > 0.97$ as the criterion for an acceptable fit. Four sets of acceptable data were used in determining each association constant. All lanes from a gel were used unless a visual inspection revealed a data point to be obviously flawed relative to neighboring points. The data were normalized using the following equation:

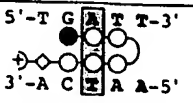
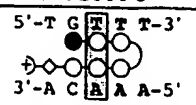
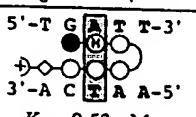
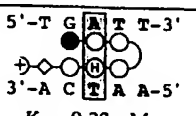
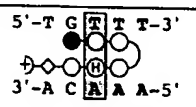
$$\theta_{norm} = \frac{\theta_{app} - \theta_{min}}{\theta_{max} - \theta_{min}} \quad (3)$$

TABLE 6 Discrimination of 5′-TGTAA-3′ and 5′-TGTTA-3′*

Pair†	5′-TGTAA-3′	5′-TGTTA-3′	K_{rel}‡
Py/Py	 K_d = 0.014 μM	 K_d = 0.007 μM	2.0
Py/Hp	 K_d = 0.20 μM	 K_d = 0.56 μM	0.36
Hp/Py	 K_d = 4.0 μM	 K_d = 0.28 μM	14

*The reported equilibrium dissociation constants are the mean values obtained from two DNase I footprint titration experiments on a 3′ ^{32}P labeled 370-bp pDEH1 *Eco*RI/*Pvu*II DNA restriction fragment[13]. The assays were carried out at 22 °C, pH 7.0 in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.
†Ring pairing opposite T•A and A•T in the third position.
‡Calculated as K_d(5′-TGTAA-3′)/K_d(5′-TGTTA-3′).

TABLE 7 Discrimination of 5'-TGTTT-3' and 5'-TGATT-3'*

Pair†	5'-TGATT-3'	5'-TGTTT-3'	K_{rel}‡
Py/Py	 K_d = 0.026 μM	 K_d = 0.005 μM	5.2
Hp/Py	 K_d = 0.53 μM	K_d = 0.008 μM	66
Py/Hp	 K_d = 0.33 μM	 K_d = 0.59 μM	0.56

*The reported equilibrium dissociation constants are the mean values obtained from two DNase I footprint titration experiments. The assays were carried out at 22 °C, pH 7.0 in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.
†Ring pairing opposite T•A and A•T in the third position.
‡Calculated as K_d(5'-TGATT -3')/K_d(5'-TGTTT-3').

EXAMPLE 5:

PREPARATION OF A BIFUNCTIONAL Hp-Py-Im-POLYAMIDE.

ImImOpPy-γ-ImPyPyPy-β-Dp-NH2. ImImOpPy-γ-ImPyPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.66 mmol/g). Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. 3-hydroxypyrrole-Boc-amino acid (0.7 mmol) was incorporated by placing the amino acid (0.5 mmol) and HBTU (0.5 mmol) in a machine synthesis cartridge. Upon automated delivery of DMF (2 mL) and DIEA (1 mL) activation occurs. A sample of ImImOpPy-γ-ImPyPyPy-β-Pam-Resin (400 mg, 0.40 mmol/gram) was placed in a glass 20 mL peptide synthesis vessel and treated with neat 3,3'-diamino-*N*-methyldipropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImImOpPy-γ-ImPyPyPy-β-Dp-NH_2 is recovered upon lyophilization of the appropriate fractions as a white powder (93 mg, 46% recovery). UV (H_2O) λ_{max} 246, 316 (66,000); ^{1}H NMR (DMSO-*d6*) δ 10.34 (s, 1 H), 10.30 (br s, 1 H), 10.25 (s, 1 H), 9.96 (s, 1 H), 9.95 (s, 1 H), 9.89 (s, 1 H), 9.24 (s, 1 H), 9.11 (s, 1 H), 8.08 (t, 1 H, *J* = 5.6 Hz), 8.0 (m, 5 H), 7.62 (s, 1 H), 7.53 (s, 1 H), 7.42 (s, 1 H), 7.23 (d, 1H, *J* = 1.2 Hz), 7.21 (m, 2 H), 7.15 (m, 2 H), 7.13 (d, 1 H), 7.11 (m, 2 H), 7.04 (d, 1 H), 6.84 (m, 3 H), 3.98 (s, 3 H), 3.97 (s, 3 H), 3.92 (s, 3 H), 3.82 (m, 6 H), 3.80 (s, 3 H), 3.77 (d, 6 H), 3.35 (q, 2 H, *J* = 5.8 Hz) 3.0-3.3 (m, 8 H), 2.86 (q, 2 H, *J* = 5.4 Hz), 2.66 (d, 3 H, *J* = 4.5 Hz), 2.31 (m, 4 H), 1.94

(quintet, 2 H, J = 6.2 Hz), 1.74 (m, 4 H); MALDI-TOF-MS (monoisotopic), 1296.0 (1296.6 calc. for $C_{60}H_{78}N_{23}O_{11}$).

ImImOpPy-γ-ImPyPyPy-β-Dp-EDTA. Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImImOpPy-γ-ImPyPyPy-β-NH_2 (13 mg, 10 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min.) and the remaining EDTA-anhydride hydrolyzed (0.1M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1% wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide ImImOpPy-γ-ImPyPyPy-β-Dp-EDTA as a white powder upon lyophilization of the appropriate fractions (5.5 mg, 40% recovery). MALDI-TOF-MS (monoisotopic), 1570.9 (1570.7 calc. for $C_{70}H_{92}N_{25}O_{18}$).

ImImHpPy-γ-ImPyPyPy-β-Dp-EDTA. In order to remove the methoxy protecting group, a sample of ImImOpPy-γ-ImPyPyPy-β-Dp-EDTA (5 mg, 3.1 μmol) was treated with sodium thiophenoxide at 100 °C for 2 h. DMF (1000 μL) and thiophenol (500 μL) were placed in a (13 x 100 mm) disposable Pyrex screw cap culture tube. A 60 % dispersion of sodium hydride in mineral oil (100 mg) was *slowly* added. Upon completion of the addition of the sodium hydride, ImImOpPy-γ-ImPyPyPy-β-Dp-EDTA (5 mg) dissolved in DMF (500 μL) was added. The solution was agitated, and placed in a 100 °C heat block, and deprotected for 2 h. Upon completion of the reaction the culture tube was cooled to 0°C, and 7 ml of a 20 % (wt/v) solution of trifluoroacetic acid added. The aqueous layer is separated from the resulting biphasic solution and purified by reversed phase HPLC. ImImHpPy-γ-ImPyPyPy-β-Dp-EDTA is recovered as a white powder upon lyophilization of the appropriate fractions (3.2 mg, 72 % recovery). UV (H_2O) λ_{max} 246, 312 (66,000); MALDI-TOF-MS (monoisotopic), 1556.6 (1556.7 calc. for $C_{69}H_{90}N_{25}O_{18}$).

ImImPyPy-γ-ImOpPyPy-β-Dp-NH_2. ImImPyPy-γ-ImOpPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.66 mmol/g). Baird, E. E. & Dervan, P. B. describes the solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996); *also see* PCT US 97/003332. 3-hydroxypyrrole-Boc-amino acid (0.7 mmol) was incorporated by placing the amino acid (0.5 mmol) and HBTU (0.5 mmol) in a machine synthesis cartridge. Upon automated delivery of DMF (2 mL) and DIEA (1 mL) activation occurs. A sample of ImImPyPy-γ-ImOpPyPy-β-Pam-Resin (400 mg, 0.40 mmol/gram) was placed in a glass 20 mL peptide synthesis vessel and treated with neat 3,3'-diamino-*N*-methyldipropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting

solution purified by reversed phase HPLC. ImImPyPy-γ-ImOpPyPy-β-Dp-NH_2 is recovered upon lyophilization of the appropriate fractions as a white powder (104 mg, 54% recovery). UV (H_2O) λ_{max} 246, 316 (66,000); MALDI-TOF-MS (monoisotopic), 1296.6 (1296.6 calc. for $C_{60}H_{78}N_{23}O_{11}$).

ImImPyPy-γ-ImOpPyPy-β-Dp-EDTA. Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImImPyPy-γ-ImOpPyPy-β-NH_2 (13 mg, 10 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min.) and the remaining EDTA-anhydride hydrolyzed (0.1M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1% wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide ImImPyPy-γ-ImOpPyPy-β-Dp-EDTA as a white powder upon lyophilization of the appropriate fractions (5.9 mg, 42% recovery). MALDI-TOF-MS (monoisotopic), 1570.8 (1570.7 calc. for $C_{70}H_{92}N_{25}O_{18}$).

ImImPyPy-γ-ImHpPyPy-β-Dp-EDTA. In order to remove the methoxy protecting group, a sample of ImImPyPy-γ-ImOpPyPy-β-Dp-EDTA (5 mg, 3.1 μmol) was treated with sodium thiophenoxide at 100 °C for 2 h. DMF (1000 μL) and thiophenol (500 μL) were placed in a (13 x 100 mm) disposable Pyrex screw cap culture tube. A 60 % dispersion of sodium hydride in mineral oil (100 mg) was *slowly* added. Upon completion of the addition of the sodium hydride, ImImPyPy-γ-ImOpPyPy-β-Dp-EDTA (5 mg) dissolved in DMF (500 μL) was added. The solution was agitated, and placed in a 100 °C heat block, and deprotected for 2 h. Upon completion of the reaction the culture tube was cooled to 0°C, and 7 ml of a 20 % (wt/v) solution of trifluoroacetic acid added. The aqueous layer is separated from the resulting biphasic solution and purified by reversed phase HPLC. ImImPyPy-γ-ImHpPyPy-β-Dp-EDTA is recovered as a white powder upon lyophilization of the appropriate fractions (3.2 mg, 72 % recovery). UV (H_2O) λ_{max} 246, 312 (66,000); MALDI-TOF-MS (monoisotopic), 1555.9 (1556.7 calc. for $C_{69}H_{90}N_{25}O_{18}$).

EXAMPLE 6:
DETERMINATION OF POLYAMIDE BINDING ORIENTATION

Affinity cleavage experiments using hairpin polyamides modified with EDTA•Fe(II) at either the C-terminus or on the γ-turn, were used to determine polyamide binding orientation and stoichiometry. The results of affinity cleavage experiments are consistent only with recognition of 6-bp by an 8-ring hairpin complex and rule out any extended 1:1 or overlapped complex formation. In addition, affinity cleavage experiments reveal hairpin formation

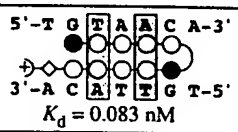
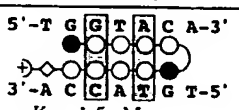
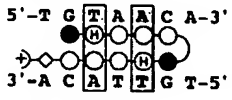
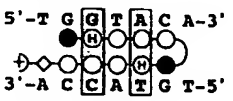
supporting the claim that it is the Hp/Py and Py/Hp pairing which form at both match and mismatch sites to discriminate A•T from T•A.

Affinity cleavage reactions were executed in a total volume of 40 μL. A stock solution of polyamide or H_2O was added to a solution containing labeled restriction fragment (20,000 cpm), affording final solution conditions of 25 mM Tris-Acetate, 20 mM NaCl, 100 μM/bp calf thymus DNA, and pH 7.0. Solutions were incubated for a minimum of 4 hours at 22°C. Subsequently, 4 μL of freshly prepared 100 μM $Fe(NH_4)_2(SO_4)_2$ was added and the solution allowed to equilibrate for 20 min. Cleavage reactions were initiated by the addition of 4 μL of 100 mM dithiothreitol, allowed to proceed for 30 min at 22 °C, then stopped by the addition of 10 μL of a solution containing 1.5 M NaOAc (pH 5.5), 0.28 mg/mL glycogen, and 14 μM base pairs calf thymus DNA, and ethanol precipitated. The reactions were resuspended in 1x TBE/80% formamide loading buffer, denatured by heating at 85 °C for 15 min, and placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-link, 7 M urea) in 1x TBE at 2000 V for 1.5 hours. Gels were dried and exposed to a storage phosphor screen. Relative cleavage intensities were determined by volume integration of individual cleavage bands using ImageQuant software.

EXAMPLE 7:
IMPROVEMENT TO POLYAMIDE SEQUENCE SPECIFICITY.

The polyamides of this invention provide improved specificity relative to existing polyamide technology. Turner, J. T., Baird, E. E., and Dervan, P.B. describe the recognition of seven base pair sequences in the minor groove of DNA by ten-ring pyrrole-imidazole polyamide hairpins *J. Am. Chem. Soc.* **1997** *119*, 7636. For example, quantitative DNaseI footprint titrations reveal that the 10-ring hairpin ImPyPyPyPy-γ-ImPyPyPyPyPy-β-Dp binds a 5'-TGTAACA-3- sequence with an equlibrium dissociation constant of 0.083 nM, and 18-fold specificity versus a single base mismatch site. A number of other sites are also bound on the 252-bp DNA fragment used for the footprint titration experiments. (Figure 13). Introduction of a Hp/Py and Py/Hp pair in the 10-ring polyamide, ImHpPyPyPy-γ-ImHpPyPyPy-β-Dp, to recognize a T•A and A•T within the 7-bp target sequence, increases the sequence-specificty. For example, a single base mismatch site 5'-TGGAACA-3 is discriminated by > 5000-fold (Figure 13, Table 8). In fact all 245 7-bp mismatch sites present on the restriction fragment are discriminated > 5000-fold by the polyamide ImHpPyPyPy-γ-ImHpPyPyPy-β-Dp (Figure 13). For cases where three A,T base pairs are present in succession it is preferred to substitute Py/Py in place of at least one Hp/Py or Py/Hp to provide for recognition of A•T and T•A at a single position.

TABLE 8 Equilibrium dissociation constants*

Polyamide†	5'-TGGTCA-3'	5'-TGGACA-3'	K_{rel}‡
Py/Py	 K_d = 0.083 nM	 K_d = 1.5 nM	18
Hp/Py	 K_d = 0.2 nM	 K_d > 1000 nM	>5000

*The reported dissociation constants are the average values obtained from three DNase I footprint titration experiments. The standard deviation for each data set is less than 15% of the reported number. Assays were carried out in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM MgCl2, and 5 mM CaCl2 at pH 7.0 and 22 °C.
†Ring pairing opposite T•A and A•T in the fourth position.
‡Calculated as K_d(5'-TGGTACA-3')/K_d(5'-TGTAACA-3').

EXAMPLE 8:
USE OF PAIRING CODE

There are 256 possible four base pair combinations of A, T, G, and C. Of these, there are a possible 240 four base pair sequences which contain at least 1 A•T or T•A base pair and

therefore can advantageously use an Hp/Py, or Py/Hp carboxamide binding. Polyamides binding to any of these sequences can be designed in accordance with the code of TABLE 2. Table 9 lists the sixteen eight-ring hairpin polyamides (1-16) which recognize the sixteen 5'-WGTNNW-3' sequences (W = A or T, X = A, G, C, or T). Table 10 lists the sixteen eight-ring hairpin polyamides (17-32) which recognize the sixteen 5'-WGANNW-3' sequences (17-32). Table 11 lists the twelve eight-ring hairpin polyamides (33-44) which recognize twelve 5'-WGGNNW-3' sequences which contain at least one A,T base pair. Table 11 lists the four eight-ring hairpin polyamides (G1-G4) which target the four 5'-WGGNNW-3' sequences (G1-G4) which contain exclusively G•C base pairs. Table 12 lists the twelve eight-ring hairpin polyamides (45-56) which recognize twelve 5'-WGCNNW-3' sequences which contain at least one A,T base pair. Table 12 lists the four eight-ring hairpin polyamides (G5-G8) which target the four 5'-WGCNNW-3' sequences (G5-G8) which contain exclusively G•C base pairs. Table 13 lists the sixteen eight-ring hairpin polyamides (57-72) which recognize the sixteen 5'-WTTNNW-3' sequences (57-72). Table 14 lists the sixteen eight-ring hairpin polyamides (73-88) which recognize the sixteen 5'-WTANNW-3' sequences (73-88). Table 15 lists the sixteen eight-ring hairpin polyamides (89-104) which recognize the sixteen 5'-WTGNNW-3' sequences (89-104). Table 16 lists the sixteen eight-ring hairpin polyamides (105-120) which recognize the sixteen 5'-WTCNNW-3' sequences (105-120). Table 17 lists the sixteen eight-ring hairpin polyamides (121-136) which recognize the sixteen 5'-WATNNW-3' sequences (121-136). Table 18 lists the sixteen eight-ring hairpin polyamides (137-152) which recognize the sixteen 5'-WAANNW-3' sequences (137-152). Table 19 lists the sixteen eight-ring hairpin polyamides (153-168) which recognize the sixteen 5'-WAGNNW-3' sequences (153-168). Table 20 lists the sixteen eight-ring hairpin polyamides (169-184) which recognize the sixteen 5'-WACNNW-3' sequences (169-184). Table 21 lists the sixteen eight-ring hairpin polyamides (185-200) which recognize the sixteen 5'-WCTNNW-3' sequences (185-200). Table 22 lists the sixteen eight-ring hairpin polyamides (201-216) which recognize the sixteen 5'-WCANNW-3' sequences (201-216). Table 23 lists the twelve eight-ring hairpin polyamides (217-228) which recognize the twelve 5'-WCGNNW-3' sequences which contain at least one A,T base pair. Table 23 lists the four eight-ring hairpin polyamides (G9-G12) which target the four 5'-WCGNNW-3' sequences (G9-G12) which contain exclusively C•G base pairs. Table 24 lists the twelve eight-ring hairpin polyamides (229-240) which recognize the twelve 5'-WCCNNW-3' sequences which contain at least one A,T base pair. Table 24 lists the four eight-ring hairpin polyamides (G13-G16) which target the four 5'-WCCNNW-3' sequences (G13-G16) which contain exclusively C•G base pairs.

TABLE 9: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WGTNNW-3'

DNA sequence	aromatic amino acid sequence
1) 5'-W G T T T W-3'	1) ImHpHpHp-γ-PyPyPyPy
2) 5'-W G T T A W-3'	2) ImHpHpPy-γ-HpPyPyPy
3) 5'-W G T T G W-3'	3) ImHpHpIm-γ-PyPyPyPy
4) 5'-W G T T C W-3'	4) ImHpHpPy-γ-ImPyPyPy
5) 5'-W G T A T W-3'	5) ImHpPyHp-γ-PyHpPyPy
6) 5'-W G T A A W-3'	6) ImHpPyPy-γ-HpHpPyPy
7) 5'-W G T A G W-3'	7) ImHpPyIm-γ-PyHpPyPy
8) 5'-W G T A C W-3'	8) ImHpPyPy-γ-ImHpPyPy
9) 5'-W G T G T W-3'	9) ImHpImHp-γ-PyPyPyPy
10) 5'-W G T G A W-3'	10) ImHpImPy-γ-HpPyPyPy
11) 5'-W G T G G W-3'	11) ImHpImIm-γ-PyPyPyPy
12) 5'-W G T G C W-3'	12) ImHpImPy-γ-ImPyPyPy
13) 5'-W G T C T W-3'	13) ImHpPyHp-γ-PyImPyPy
14) 5'-W G T C A W-3'	14) ImHpPyPy-γ-HpImPyPy
15) 5'-W G T C G W-3'	15) ImHpPyIm-γ-PyImPyPy
16) 5'-W G T C C W-3'	16) ImHpPyPy-γ-ImImPyPy

TABLE 10: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WGANNW-3'

DNA sequence	aromatic amino acid sequence
17) 5'-W G A T T W-3'	17) ImPyHpHp-γ-PyPyHpPy
18) 5'-W G A T A W-3'	18) ImPyHpPy-γ-HpPyHpPy
19) 5'-W G A T G W-3'	19) ImPyHpIm-γ-PyPyHpPy
20) 5'-W G A T C W-3'	20) ImPyHpPy-γ-ImPyHpPy
21) 5'-W G A A T W-3'	21) ImPyPyHp-γ-PyHpHpPy
22) 5'-W G A A A W-3'	22) ImPyPyPy-γ-HpHpHpPy
23) 5'-W G A A G W-3'	23) ImPyPyIm-γ-PyHpHpPy
24) 5'-W G A A C W-3'	24) ImPyPyPy-γ-ImHpHpPy
25) 5'-W G A G T W-3'	25) ImPyImHp-γ-PyPyHpPy
26) 5'-W G A G A W-3'	26) ImPyImPy-γ-HpPyHpPy
27) 5'-W G A G G W-3'	27) ImPyImIm-γ-PyPyHpPy
28) 5'-W G A G C W-3'	28) ImPyImPy-γ-ImPyHpPy
29) 5'-W G A C T W-3'	29) ImPyPyHp-γ-PyImHpPy
30) 5'-W G A C A W-3'	30) ImPyPyPy-γ-HpImHpPy
31) 5'-W G A C G W-3'	31) ImPyPyIm-γ-PyImHpPy
32) 5'-W G A C C W-3'	32) ImPyPyPy-γ-ImImHpPy

TABLE 11: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WGGNNW-3'

DNA sequence	aromatic amino acid sequence
33) 5'-W G G T T W-3'	33) ImImHpHp-γ-PyPyPyPy
34) 5'-W G G T A W-3'	34) ImImHpPy-γ-HpPyPyPy
35) 5'-W G G T G W-3'	35) ImImHpIm-γ-PyPyPyPy
36) 5'-W G G T C W-3'	36) ImImHpPy-γ-ImPyPyPy
37) 5'-W G G A T W-3'	37) ImImPyHp-γ-PyHpPyPy
38) 5'-W G G A A W-3'	38) ImImPyPy-γ-HpHpPyPy
39) 5'-W G G A G W-3'	39) ImImPyIm-γ-PyHpPyPy
40) 5'-W G G A C W-3'	40) ImImPyPy-γ-ImHpPyPy
41) 5'-W G G G T W-3'	41) ImImImHp-γ-PyPyPyPy
42) 5'-W G G G A W-3'	42) ImImImPy-γ-HpPyPyPy
43) 5'-W G G C T W-3'	43) ImImPyHp-γ-PyImPyPy
44) 5'-W G G C A W-3'	44) ImImPyPy-γ-HpImPyPy
G1) 5'-W G G G G W-3'	G1) ImImImIm-γ-PyPyPyPy
G2) 5'-W G G G C W-3'	G2) ImImImPy-γ-ImPyPyPy
G3) 5'-W G G C G W-3'	G3) ImImPyIm-γ-PyImPyPy
G4) 5'-W G G C C W-3'	G4) ImImPyPy-γ-ImImPyPy

TABLE 12: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WGCNNW-3'

DNA sequence	aromatic amino acid sequence
45) 5'-W G C T T W-3'	45) ImPyHpHp-γ-PyPyImPy
46) 5'-W G C T A W-3'	46) ImPyHpPy-γ-HpPyImPy
47) 5'-W G C T G W-3'	47) ImPyHpIm-γ-PyPyImPy
48) 5'-W G C T C W-3'	48) ImPyHpPy-γ-ImPyImPy
49) 5'-W G C A T W-3'	49) ImPyPyHp-γ-PyHpImPy
50) 5'-W G C A A W-3'	50) ImPyPyPy-γ-HpHpImPy
51) 5'-W G C A G W-3'	51) ImPyPyIm-γ-PyHpImPy
52) 5'-W G C A C W-3'	52) ImPyPyPy-γ-ImHpImPy
53) 5'-W G C G T W-3'	53) ImPyImHp-γ-PyPyImPy
54) 5'-W G C G A W-3'	54) ImPyImPy-γ-HpPyImPy
55) 5'-W G C C T W-3'	55) ImPyPyHp-γ-PyImImPy
56) 5'-W G C C A W-3'	56) ImPyPyPy-γ-HpImImPy
G5) 5'-W G C G G W-3'	G5) ImPyImIm-γ-PyPyImPy
G6) 5'-W G C G C W-3'	G6) ImPyImPy-γ-ImPyImPy
G7) 5'-W G C C G W-3'	G7) ImPyPyIm-γ-PyImImPy
G8) 5'-W G C C C W-3'	G8) ImPyPyPy-γ-ImImImPy

TABLE 13: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WTTNNW-3'

DNA sequence	aromatic amino acid sequence
57) 5'-W T T T T W-3'	57) HpHpHpHp-γ-PyPyPyPy
58) 5'-W T T T A W-3'	58) HpHpHpPy-γ-HpPyPyPy
59) 5'-W T T T G W-3'	59) HpHpHpIm-γ-PyPyPyPy
60) 5'-W T T T C W-3'	60) HpHpHpPy-γ-ImPyPyPy
61) 5'-W T T A T W-3'	61) HpHpPyHp-γ-PyHpPyPy
62) 5'-W T T A A W-3'	62) HpHpPyPy-γ-HpHpPyPy
63) 5'-W T T A G W-3'	63) HpHpPyIm-γ-PyHpPyPy
64) 5'-W T T A C W-3'	64) HpHpPyPy-γ-ImHpPyPy
65) 5'-W T T G T W-3'	65) HpHpImHp-γ-PyPyPyPy
66) 5'-W T T G A W-3'	66) HpHpImPy-γ-HpPyPyPy
67) 5'-W T T G G W-3'	67) HpHpImIm-γ-PyPyPyPy
68) 5'-W T T G C W-3'	68) HpHpImPy-γ-ImPyPyPy
69) 5'-W T T C T W-3'	69) HpHpPyHp-γ-PyImPyPy
70) 5'-W T T C A W-3'	70) HpHpPyPy-γ-HpImPyPy
71) 5'-W T T C G W-3'	71) HpHpPyIm-γ-PyImPyPy
72) 5'-W T T C C W-3'	72) HpHpPyPy-γ-ImImPyPy

TABLE 14: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WTANNW-3'

DNA sequence	aromatic amino acid sequence
73) 5'-W T A T T W-3'	73)HpPyHpHp-γ-PyPyHpPy
74) 5'-W T A T A W-3'	74)HpPyHpPy-γ-HpPyHpPy
75) 5'-W T A T G W-3'	75)HpPyHpIm-γ-PyPyHpPy
76) 5'-W T A T C W-3'	76)HpPyHpPy-γ-ImPyHpPy
77) 5'-W T A A T W-3'	77)HpPyPyHp-γ-PyHpHpPy
78) 5'-W T A A A W-3'	78)HpPyPyPy-γ-HpHpHpPy
79) 5'-W T A A G W-3'	79)HpPyPyIm-γ-PyHpHpPy
80) 5'-W T A A C W-3'	80)HpPyPyPy-γ-ImHpHpPy
81) 5'-W T A G T W-3'	81)HpPyImHp-γ-PyPyHpPy
82) 5'-W T A G A W-3'	82)HpPyImPy-γ-HpPyHpPy
83) 5'-W T A G G W-3'	83)HpPyImIm-γ-PyPyHpPy
84) 5'-W T A G C W-3'	84)HpPyImPy-γ-ImPyHpPy
85) 5'-W T A C T W-3'	85)HpPyPyHp-γ-PyImHpPy
86) 5'-W T A C A W-3'	86)HpPyPyPy-γ-HpImHpPy
87) 5'-W T A C G W-3'	87)HpPyPyIm-γ-PyImHpPy
88) 5'-W T A C C W-3'	88)HpPyPyPy-γ-ImImHpPy

TABLE 15: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WTGNNW-3'

	DNA sequence	aromatic amino acid sequence
89)	5'-W T G T T W-3'	89) HpImHpHp-γ-PyPyPyPy
90)	5'-W T G T A W-3'	90) HpImHpPy-γ-HpPyPyPy
91)	5'-W T G T G W-3'	91) HpImHpIm-γ-PyPyPyPy
92)	5'-W T G T C W-3'	92) HpImHpPy-γ-ImPyPyPy
93)	5'-W T G A T W-3'	93) HpImPyHp-γ-PyHpPyPy
94)	5'-W T G A A W-3'	94) HpImPyPy-γ-HpHpPyPy
95)	5'-W T G A G W-3'	95) HpImPyIm-γ-PyHpPyPy
96)	5'-W T G A C W-3'	96) HpImPyPy-γ-ImHpPyPy
97)	5'-W T G G T W-3'	97) HpImImHp-γ-PyPyPyPy
98)	5'-W T G G A W-3'	98) HpImImPy-γ-HpPyPyPy
99)	5'-W T G C T W-3'	99) HpImPyHp-γ-PyImPyPy
100)	5'-W T G C A W-3'	100) HpImPyPy-γ-HpImPyPy
101)	5'-W T G G G W-3'	101) HpImImIm-γ-PyPyPyPy
102)	5'-W T G G C W-3'	102) HpImImPy-γ-ImPyPyPy
103)	5'-W T G C G W-3'	103) HpImPyIm-γ-PyImPyPy
104)	5'-W T G C C W-3'	104) HpImPyPy-γ-ImImPyPy

TABLE 16: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WTCNNW-3'

DNA sequence	aromatic amino acid sequence
105) 5'-W T C T T W-3'	105) HpPyHpHp-γ-PyPyImPy
106) 5'-W T C T A W-3'	106) HpPyHpPy-γ-HpPyImPy
107) 5'-W T C T G W-3'	107) HpPyHpIm-γ-PyPyImPy
108) 5'-W T C T C W-3'	108) HpPyHpPy-γ-ImPyImPy
109) 5'-W T C A T W-3'	109) HpPyPyHp-γ-PyHpImPy
110) 5'-W T C A A W-3'	110) HpPyPyPy-γ-HpHpImPy
111) 5'-W T C A G W-3'	111) HpPyPyIm-γ-PyHpImPy
112) 5'-W T C A C W-3'	112) HpPyPyPy-γ-ImHpImPy
113) 5'-W T C G T W-3'	113) HpPyImHp-γ-PyPyImPy
114) 5'-W T C G A W-3'	114) HpPyImPy-γ-HpPyImPy
115) 5'-W T C C T W-3'	115) HpPyPyHp-γ-PyImImPy
116) 5'-W T C C A W-3'	116) HpPyPyPy-γ-HpImImPy
117) 5'-W T C G G W-3'	117) HpPyImIm-γ-PyPyImPy
118) 5'-W T C G C W-3'	118) HpPyImPy-γ-ImPyImPy
119) 5'-W T C C G W-3'	119) HpPyPyIm-γ-PyImImPy
120) 5'-W T C C C W-3'	120) HpPyPyPy-γ-ImImImPy

TABLE 17: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WATNNW-3'

DNA sequence	aromatic amino acid sequence
121) 5'-W A T T T W-3'	121) PyHpHpHp-γ-PyPyPyHp
122) 5'-W A T T A W-3'	122) PyHpHpPy-γ-HpPyPyHp
123) 5'-W A T T G W-3'	123) PyHpHpIm-γ-PyPyPyHp
124) 5'-W A T T C W-3'	124) PyHpHpPy-γ-ImPyPyHp
125) 5'-W A T A T W-3'	125) PyHpPyHp-γ-PyHpPyHp
126) 5'-W A T A A W-3'	126) PyHpPyPy-γ-HpHpPyHp
127) 5'-W A T A G W-3'	127) PyHpPyIm-γ-PyHpPyHp
128) 5'-W A T A C W-3'	128) PyHpPyPy-γ-ImHpPyHp
129) 5'-W A T G T W-3'	129) PyHpImHp-γ-PyPyPyHp
130) 5'-W A T G A W-3'	130) PyHpImPy-γ-HpPyPyHp
131) 5'-W A T G G W-3'	131) PyHpImIm-γ-PyPyPyHp
132) 5'-W A T G C W-3'	132) PyHpImPy-γ-ImPyPyHp
133) 5'-W A T C T W-3'	133) PyHpPyHp-γ-PyImPyHp
134) 5'-W A T C A W-3'	134) PyHpPyPy-γ-HpImPyHp
135) 5'-W A T C G W-3'	135) PyHpPyIm-γ-PyImPyHp
136) 5'-W A T C C W-3'	136) PyHpPyPy-γ-ImImPyHp

TABLE 18: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WAANNW-3'

DNA sequence	aromatic amino acid sequence
137) 5'-W A A T T W-3'	137) PyPyHpHp-γ-PyPyHpHp
138) 5'-W A A T A W-3'	138) PyPyHpPy-γ-HpPyHpHp
139) 5'-W A A T G W-3'	139) PyPyHpIm-γ-PyPyHpHp
140) 5'-W A A T C W-3'	140) PyPyHpPy-γ-ImPyHpHp
141) 5'-W A A A T W-3'	141) PyPyPyHp-γ-PyHpHpHp
142) 5'-W A A A A W-3'	142) PyPyPyPy-γ-HpHpHpHp
143) 5'-W A A A G W-3'	143) PyPyPyIm-γ-PyHpHpHp
144) 5'-W A A A C W-3'	144) PyPyPyPy-γ-ImHpHpHp
145) 5'-W A A G T W-3'	145) PyPyImHp-γ-PyPyHpHp
146) 5'-W A A G A W-3'	146) PyPyImPy-γ-HpPyHpHp
147) 5'-W A A G G W-3'	147) PyPyImIm-γ-PyPyHpHp
148) 5'-W A A G C W-3'	148) PyPyImPy-γ-ImPyHpHp
149) 5'-W A A C T W-3'	149) PyPyPyHp-γ-PyImHpHp
150) 5'-W A A C A W-3'	150) PyPyPyPy-γ-HpImHpHp
151) 5'-W A A C G W-3'	151) PyPyPyIm-γ-PyImHpHp
152) 5'-W A A C C W-3'	152) PyPyPyPy-γ-ImImHpHp

TABLE 19: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WAGNNW-3'

DNA sequence	aromatic amino acid sequence
153) 5'-W A G T T W-3'	153) PyImHpHp-γ-PyPyPyHp
154) 5'-W A G T A W-3'	154) PyImHpPy-γ-HpPyPyHp
155) 5'-W A G T G W-3'	155) PyImHpIm-γ-PyPyPyHp
156) 5'-W A G T C W-3'	156) PyImHpPy-γ-ImPyPyHp
157) 5'-W A G A T W-3'	157) PyImPyHp-γ-PyHpPyHp
158) 5'-W A G A A W-3'	158) PyImPyPy-γ-HpHpPyHp
159) 5'-W A G A G W-3'	159) PyImPyIm-γ-PyHpPyHp
160) 5'-W A G A C W-3'	160) PyImPyPy-γ-ImHpPyHp
161) 5'-W A G G T W-3'	161) PyImImHp-γ-PyPyPyHp
162) 5'-W A G G A W-3'	162) PyImImPy-γ-HpPyPyHp
163) 5'-W A G C T W-3'	163) PyImPyHp-γ-PyImPyHp
164) 5'-W A G C A W-3'	164) PyImPyPy-γ-HpImPyHp
165) 5'-W A G G G W-3'	165) PyImImIm-γ-PyPyPyHp
166) 5'-W A G G C W-3'	166) PyImImPy-γ-ImPyPyHp
167) 5'-W A G C G W-3'	167) PyImPyIm-γ-PyImPyHp
168) 5'-W A G C C W-3'	168) PyImPyPy-γ-ImImPyHp

TABLE 20: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WACNNW-3'

DNA sequence	aromatic amino acid sequence
169) 5'-W A C T T W-3'	169) PyPyHpHp-γ-PyPyImHp
170) 5'-W A C T A W-3'	170) PyPyHpPy-γ-HpPyImHp
171) 5'-W A C T G W-3'	171) PyPyHpIm-γ-PyPyImHp
172) 5'-W A C T C W-3'	172) PyPyHpPy-γ-ImPyImHp
173) 5'-W A C A T W-3'	173) PyPyPyHp-γ-PyHpImHp
174) 5'-W A C A A W-3'	174) PyPyPyPy-γ-HpHpImHp
175) 5'-W A C A G W-3'	175) PyPyPyIm-γ-PyHpImHp
176) 5'-W A C A C W-3'	176) PyPyPyPy-γ-ImHpImHp
177) 5'-W A C G T W-3'	177) PyPyImHp-γ-PyPyImHp
178) 5'-W A C G A W-3'	178) PyPyImPy-γ-HpPyImHp
179) 5'-W A C C T W-3'	179) PyPyPyHp-γ-PyImImHp
180) 5'-W A C C A W-3'	180) PyPyPyPy-γ-HpImImHp
181) 5'-W A C G G W-3'	181) PyPyImIm-γ-PyPyImHp
182) 5'-W A C G C W-3'	182) PyPyImPy-γ-ImPyImHp
183) 5'-W A C C G W-3'	183) PyPyPyIm-γ-PyImImHp
184) 5'-W A C C C W-3'	184) PyPyPyPy-γ-ImImImHp

TABLE 21: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WCTNNW-3'

DNA sequence	aromatic amino acid sequence
185) 5'-W C T T T W-3'	185) PyHpHpHp-γ-PyPyPyIm
186) 5'-W C T T A W-3'	186) PyHpHpPy-γ-HpPyPyIm
187) 5'-W C T T G W-3'	187) PyHpHpIm-γ-PyPyPyIm
188) 5'-W C T T C W-3'	188) PyHpHpPy-γ-ImPyPyIm
189) 5'-W C T A T W-3'	189) PyHpPyHp-γ-PyHpPyIm
190) 5'-W C T A A W-3'	190) PyHpPyPy-γ-HpHpPyIm
191) 5'-W C T A G W-3'	191) PyHpPyIm-γ-PyHpPyIm
192) 5'-W C T A C W-3'	192) PyHpPyPy-γ-ImHpPyIm
193) 5'-W C T G T W-3'	193) PyHpImHp-γ-PyPyPyIm
194) 5'-W C T G A W-3'	194) PyHpImPy-γ-HpPyPyIm
195) 5'-W C T G G W-3'	195) PyHpImIm-γ-PyPyPyIm
196) 5'-W C T G C W-3'	196) PyHpImPy-γ-ImPyPyIm
197) 5'-W C T C T W-3'	197) PyHpPyHp-γ-PyImPyIm
198) 5'-W C T C A W-3'	198) PyHpPyPy-γ-HpImPyIm
199) 5'-W C T C G W-3'	199) PyHpPyIm-γ-PyImPyIm
200) 5'-W C T C C W-3'	200) PyHpPyPy-γ-ImImPyIm

TABLE 22: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WCANNW-3'

DNA sequence	aromatic amino acid sequence
201) 5'-W C A T T W-3'	201) PyPyHpHp-γ-PyPyHpIm
202) 5'-W C A T A W-3'	202) PyPyHpPy-γ-HpPyHpIm
203) 5'-W C A T G W-3'	203) PyPyHpIm-γ-PyPyHpIm
204) 5'-W C A T C W-3'	204) PyPyHpPy-γ-ImPyHpIm
205) 5'-W C A A T W-3'	205) PyPyPyHp-γ-PyHpHpIm
206) 5'-W C A A A W-3'	206) PyPyPyPy-γ-HpHpHpIm
207) 5'-W C A A G W-3'	207) PyPyPyIm-γ-PyHpHpIm
208) 5'-W C A A C W-3'	208) PyPyPyPy-γ-ImHpHpIm
209) 5'-W C A G T W-3'	209) PyPyImHp-γ-PyPyHpIm
210) 5'-W C A G A W-3'	210) PyPyImPy-γ-HpPyHpIm
211) 5'-W C A G G W-3'	211) PyPyImIm-γ-PyPyHpIm
212) 5'-W C A G C W-3'	212) PyPyImPy-γ-ImPyHpIm
213) 5'-W C A C T W-3'	213) PyPyPyHp-γ-PyImHpIm
214) 5'-W C A C A W-3'	214) PyPyPyPy-γ-HpImHpIm
215) 5'-W C A C G W-3'	215) PyPyPyIm-γ-PyImHpIm
216) 5'-W C A C C W-3'	216) PyPyPyPy-γ-ImImHpIm

TABLE 23: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WCGNNW-3'

	DNA sequence	aromatic amino acid sequence
217)	5'-W C G T T W-3'	217) PyImHpHp-γ-PyPyPyIm
218)	5'-W C G T A W-3'	218) PyImHpPy-γ-HpPyPyIm
219)	5'-W C G T G W-3'	219) PyImHpIm-γ-PyPyPyIm
220)	5'-W C G T C W-3'	220) PyImHpPy-γ-ImPyPyIm
221)	5'-W C G A T W-3'	221) PyImPyHp-γ-PyHpPyIm
222)	5'-W C G A A W-3'	222) PyImPyPy-γ-HpHpPyIm
223)	5'-W C G A G W-3'	223) PyImPyIm-γ-PyHpPyIm
224)	5'-W C G A C W-3'	224) PyImPyPy-γ-ImHpPyIm
225)	5'-W C G G T W-3'	225) PyImImHp-γ-PyPyPyIm
226)	5'-W C G G A W-3'	226) PyImImPy-γ-HpPyPyIm
227)	5'-W C G C T W-3'	227) PyImPyHp-γ-PyImPyIm
228)	5'-W C G C A W-3'	228) PyImPyPy-γ-HpImPyIm
G9)	5'-W C G G G W-3'	G9) PyImImIm-γ-PyPyPyIm
G10)	5'-W C G G C W-3'	G10) PyImImPy-γ-ImPyPyIm
G11)	5'-W C G C G W-3'	G11) PyImPyIm-γ-PyImPyIm
G12)	5'-W C G C C W-3'	G12) PyImPyPy-γ-ImImPyIm

TABLE 24: 8-ring Hairpin Polyamides for recognition of 6-bp 5'-WCCNNW-3'

DNA sequence	aromatic amino acid sequence
229) 5'-W C C T T W-3'	229) PyPyHpHp-γ-PyPyImIm
230) 5'-W C-C T A W-3'	230) PyPyHpPy-γ-HpPyImIm
231) 5'-W C C T G W-3'	231) PyPyHpIm-γ-PyPyImIm
232) 5'-W C C T C W-3'	232) PyPyHpPy-γ-ImPyImIm
233) 5'-W C C A T W-3'	233) PyPyPyHp-γ-PyHpImIm
234) 5'-W C C A A W-3'	234) PyPyPyPy-γ-HpHpImIm
235) 5'-W C C A G W-3'	235) PyPyPyIm-γ-PyHpImIm
236) 5'-W C C A C W-3'	236) PyPyPyPy-γ-ImHpImIm
237) 5'-W C C G T W-3'	237) PyPyImHp-γ-PyPyImIm
238) 5'-W C C G A W-3'	238) PyPyImPy-γ-HpPyImIm
239) 5'-W C C C T W-3'	239) PyPyPyHp-γ-PyImImIm
240) 5'-W C C C A W-3'	240) PyPyPyPy-γ-HpImImIm
G13) 5'-W C C G G W-3'	G13) PyPyImIm-γ-PyPyImIm
G14) 5'-W C C G C W-3'	G14) PyPyImPy-γ-ImPyImIm
G15) 5'-W C C C G W-3'	G15) PyPyPyIm-γ-PyImImIm
G16) 5'-W C C C C W-3'	G16) PyPyPyPy-γ-ImImImIm

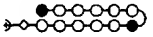
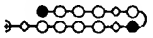
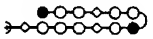
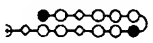
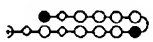
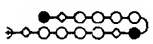
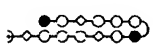
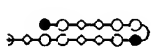
EXAMPLE 9:

Aliphatic/Aromatic amino acid pairing for recognition of the DNA minor groove.

Selective placement of an aliphatic β-alanine (β) residue paired side-by-side with either a pyrrole (Py) or imidazole (Im) aromatic amino acid is found to compensate for sequence composition effects for recognition of the minor groove of DNA by hairpin pyrrole-imidazole polyamides. A series of polyamides were prepared which contain pyrrole and imidazole aromatic amino acids, as well as γ-aminobutyric acid (γ) "turn" and β-alanine "spring" aliphatic amino acid residues. The binding affinities and specificities of these polyamides are regulated by the placement of paired β/β Py/β and Im/β residues. Quantitative footprint titrations demonstrate that replacing two Py/Py pairings in a 12-ring hairpin (6-γ-6) with two Py/β

pairings affords 10-fold enhanced affinity and similar sequence specificity for an 8-bp target sequence.

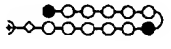
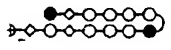
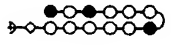
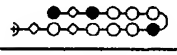
Table 25 Equilibrium association constants (M^{-1}) for polyamides.[a-c]

Polyamide	5'-TGTTAACA-3'	5'-TGTGAACA-3'	Specificity[d]
	2.5×10^9	3.9×10^8	6
	1.3×10^9	2.0×10^8	7
	1.7×10^{10}	2.7×10^9	6
	**1.2×10^{11}**	2.2×10^9	**55**
	6.6×10^9	2.5×10^8	**26**
	4.5×10^{10}	7.7×10^9	6
	2.7×10^{10}	5.7×10^9	5
	$\leq 1 \times 10^8$	$\leq 1 \times 10^8$	1

[a]Values reported are the mean values obtained from three DNase I footprint titration experiments. [b]The assays were carried out at 22 °C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. [c]Match site association constants and specificities higher than the parent hairpin are shown in boldtype. [d]Specificity is calculated as K_a(match) / K_a(mismatch).

The 6-γ-6 hairpin ImPyImPyPyPy-γ-ImPyPyPyPyPy-β-Dp, which contains six consecutive amino acid pairings, is unable to discriminate a single-base-pair mismatch site 5'-TGTTAACA-3' from a 5'-TGTGAACA-3' match site. The hairpin polyamide Im-β-ImPyPyPy-γ-ImPyPyPy-β-Py-β-Dp binds to the 8-bp match sequence 5'-TGTGAACA-3' with an equilibrium association constant of $K_a = 2.4 \times 10^{10}\ M^{-1}$ and > 48-fold specificity versus the 5'-TGTTAACA-3' single-base-pair mismatch site.

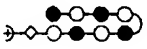
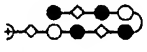
Table 26 Equilibrium association constants (M^{-1}) for polyamides.[a-c]

Polyamide	5'-TGTTAACA-3'	5'-TGTGAACA-3'	Specificity[d]
	2.5×10^9	3.9×10^8	6
	**6.6×10^9**	2.5×10^8	**26**
	5×10^9	5×10^9	1
	$\leq 5 \times 10^8$	**2.4×10^{10}**	**≥ 48**

[a] Values reported for **1**, **5**, and **10** are the mean values obtained from three DNase I footprint titration experiments. [b] The assays were carried out at 22 °C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. [c] Match site association constants and specificities higher than parent hairpins are shown in boldtype. [d] Specificity is calculated as K_a(match) / K_a(mismatch).

Modeling indicates that the β-alanine residue relaxes ligand curvature, providing for optimal hydrogen bond formation between the floor of the minor groove and both Im-residues within the Im-β-Im polyamide subunit. This observation provided the basis for design of a hairpin polyamide, Im-β-ImPy-γ-Im-β-ImPy-β-Dp, which incorporates Im/β pairings to recognize a "problematic" 5'-GCGC-3' sequence at subnanomolar concentrations.

Table 27 Equilibrium association constants (M^{-1}) for polyamides.[a-b]

Polyamide	5'-TGCGCA-3'	5'-TGGCCA-3'	5'-TGGGGA-3'
	3.7×10^7	$< 10^7$	$< 10^7$
	3.7×10^9	1.4×10^8	1.1×10^8

[a] Values reported are the mean values obtained from a minimum of three DNase I footprint titration experiments. [b] The assays were carried out at 22 °C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

These results identify Im/β and β/Im pairings that respectively discriminate G•C and C•G from A•T/T•A as well as Py/β and β/Py pairings that discriminate A•T/T•A from G•C/C•G. These aliphatic/aromatic amino acid pairings will facilitate the design of hairpin polyamides which recognize both a larger binding site size as well as a more diverse sequence repertoire.

EXAMPLE 10:

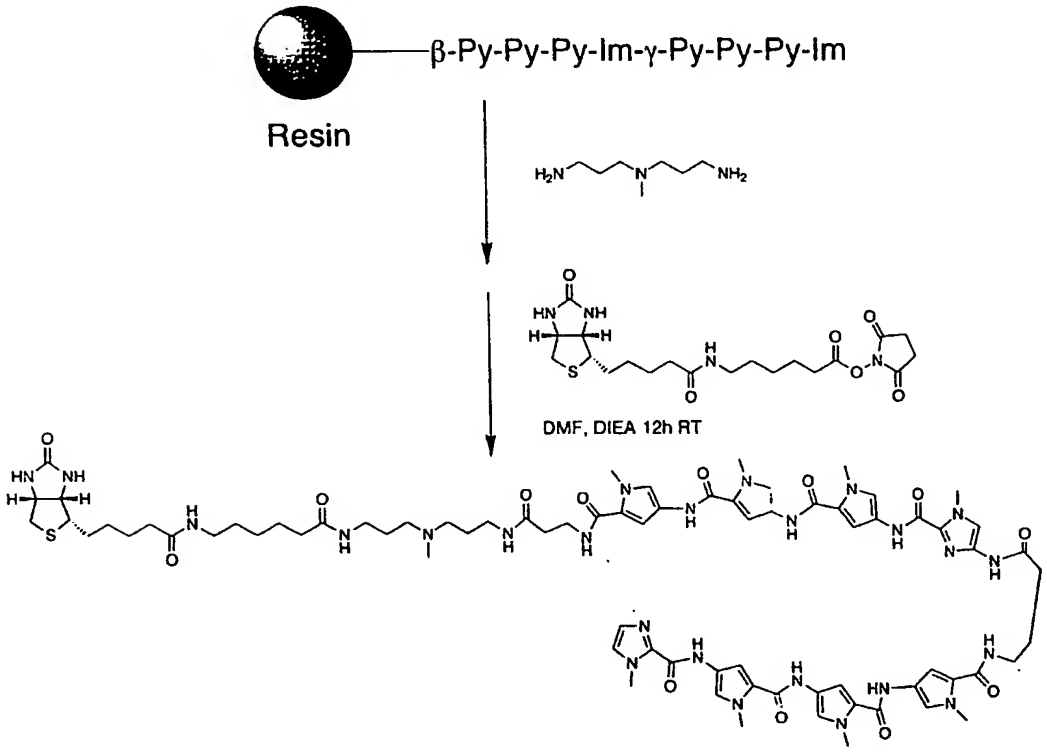
POLYAMIDE BIOTIN CONJUGATES

Bifunctional conjugates prepared between sequence specific DNA binding polyamides and biotin are useful for a variety of applications. First, such compounds can be readily attached

to a variety of matrices through the strong interaction of biotin with the protein streptavidin. Readily available strepdavidin-derivatized matrices include magnetic beads for separations as well as resins for chromatography.

A number of such polyamide-biotin conjugates have been synthesized by solid phase synthetic methods outlined in detail above. Following resin cleavage with a variety of diamines, the polyamides were reacted with various biotin carboxylic acid derivatives to yield bifunctional conjugates. The bifunctional conjugates were purified by HPLC and characterized by MALDI-TOF mass spectroscopy and ^{1}H NMR.

The scheme for the synthesis of an exemplary biotin-polyamide conjugate is shown below.



The foregoing is intended to be illustrative of the present invention, but not limiting. Numerous variations and modifications of the present invention may be effected without departing from the true spirit and scope of the invention.

What is claimed is:

1. In a polyamide having at least three consecutive carboxamide pairs for binding to at least three DNA base pairs in the minor groove of a duplex DNA sequence having at least one A•T or T•A DNA base pair, the improvement comprising selecting a Hp/Py carboxamide pair to correspond to a T•A base pair in the minor groove of the duplex DNA sequence or selecting a Py/Hp carboxamide pair to bind to an A•T DNA base pair in the minor groove of the duplex DNA sequence.

2. The polyamide of claim 1 wherein at least four consecutive carboxamide pairs bind to at least four DNA base pairs.

3. The polyamide of claim 1 wherein at least five consecutive carboxamide pairs bind to at least five DNA base pairs.

4. The polyamide of claim 1 wherein at least six consecutive carboxamide pairs bind to at least six DNA base pairs.

5. The polyamide of claim 1 wherein the A•T or T•A base pair has a G•C or C•G base pair on either side.

6. The polyamide of claim 1 wherein the duplex DNA sequence is a regulatory sequence.

7. The polyamide of claim 1 wherein the duplex DNA sequence is a promoter sequence.

8. The polyamide of claim 1 wherein the duplex DNA sequence is a coding sequence.

9. The polyamide of claim 1 wherein the duplex DNA sequence is a non-coding sequence.

10. The polyamide of claim 1 wherein the binding of the carboxamide pairs to the DNA base pairs modulates the expression of a gene.

11. A composition comprising an effective amount of the polyamide of claim 1 and a pharmologically suitable excipient.

12. A diagnostic kit comprising the polyamide of claim 1.

13. A polyamide according to claim 1 having the formula:

$$X_1X_2X_3X_4\text{-}\gamma\text{-}X_5X_6X_7X_8$$

wherein γ is -NH-CH_2-CH_2-CH_2-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R–2,4-diaminobutyric acid; X_4/X_5, X_3/X_6, X_2/X_7, and X_1/X_8 represent carboxamide binding pairs which bind the DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other binding pairs are selected from Py/Im Im/Py to correspond to the DNA base pair in the minor groove to be bound.

14. The polyamide of claim 13 wherein there is at least one β-alanine in a non- Hp containing binding pair.

15. The polyamide of claim 13 wherein dimethylaminopropylamide is covalently bound to X_1 or X_8.

16. A polyamide selected from those listed in Tables 9-24 as compounds 1 through 240.

17. A polyamide selected from shown in Fig. 4.

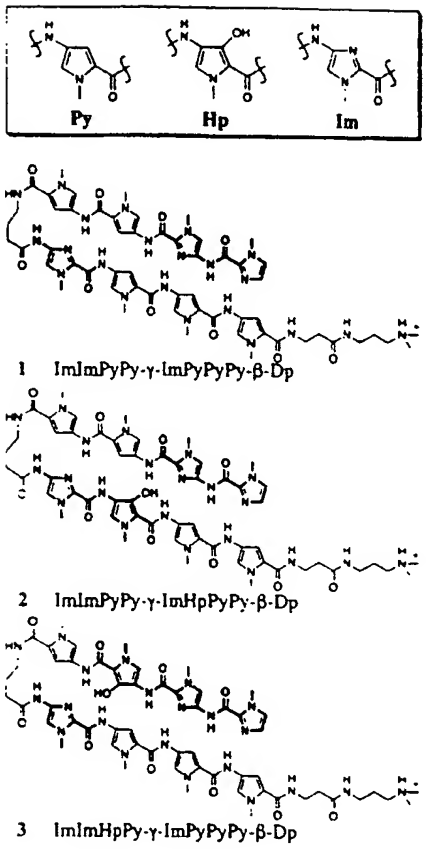

Py
Hp
Im
1 ImImPyPy-γ-ImPyPyPy-β-Dp
2 ImImPyPy-γ-ImHpPyPy-β-Dp
3 ImImHpPy-γ-ImPyPyPy-β-Dp

FIG. 1

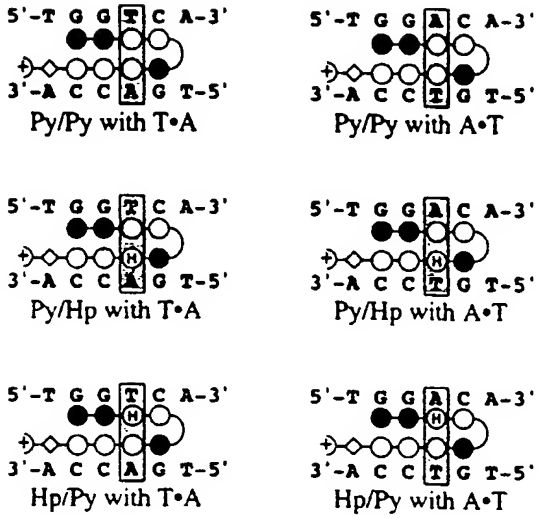

5'-T G G T C A-3'
3'-A C C A G T-5'
Py/Py with T•A
5'-T G G A C A-3'
3'-A C C T G T-5'
Py/Py with A•T
5'-T G G T C A-3'
3'-A C C A G T-5'
Py/Hp with T•A
5'-T G G A C A-3'
3'-A C C T G T-5'
Py/Hp with A•T
5'-T G G T C A-3'
3'-A C C A G T-5'
Hp/Py with T•A
5'-T G G A C A-3'
3'-A C C T G T-5'
Hp/Py with A•T

FIG. 2

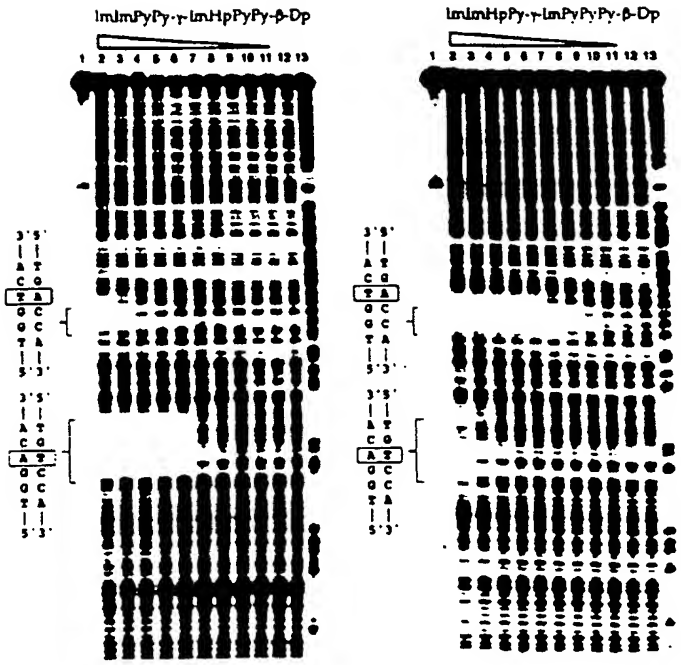
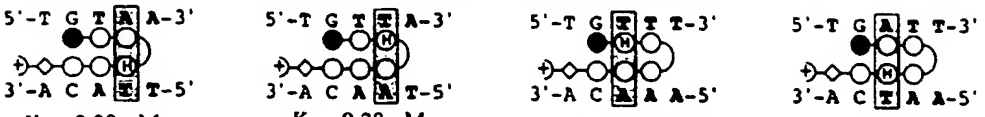
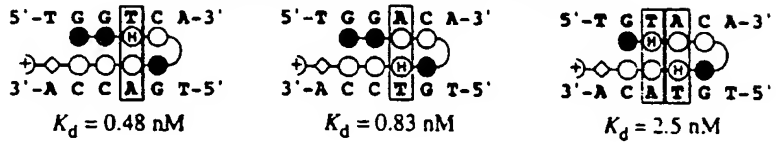


FIG. 3

6-Ring Hairpin Hp-Py-Im-Polyamides



8-Ring Hairpin Hp-Py-Im-Polyamides



10-Ring Hairpin Hp-Py-Im-Polyamides

FIG. 4

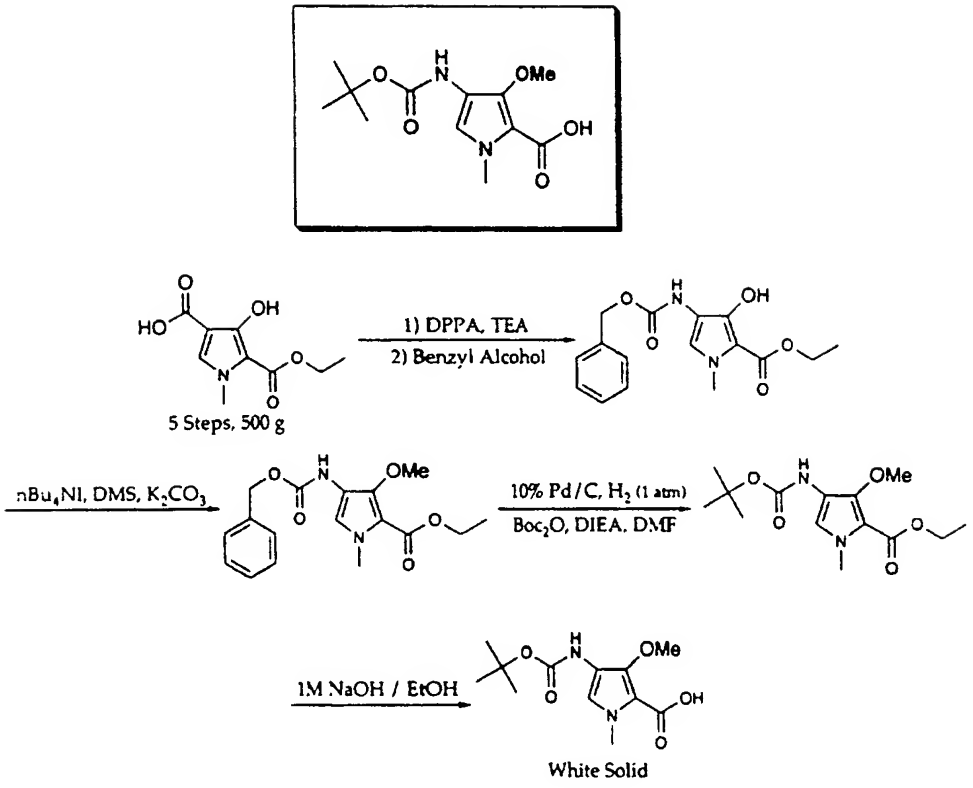
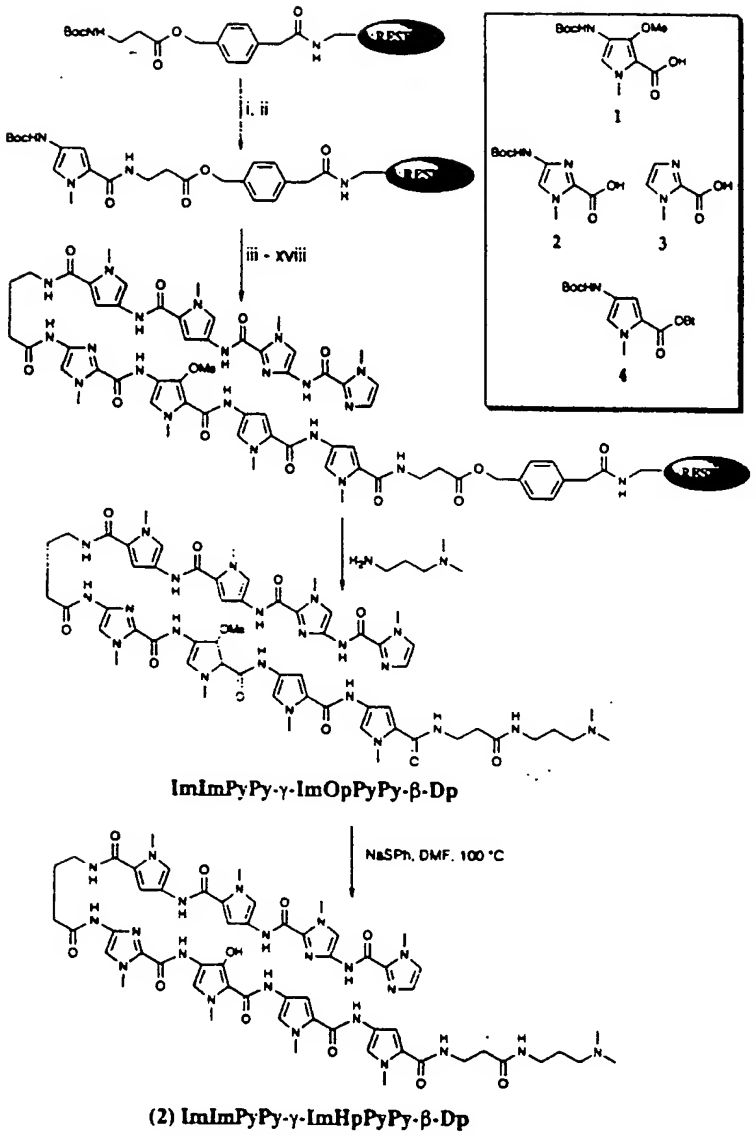

OMe
OH
1) DPPA, TEA
2) Benzyl Alcohol
5 Steps, 500 g
nBu4NI, DMS, K2CO3
10% Pd/C, H2 (1 atm)
Boc2O, DIEA, DMF
1M NaOH / EtOH
White Solid

FIG. 5

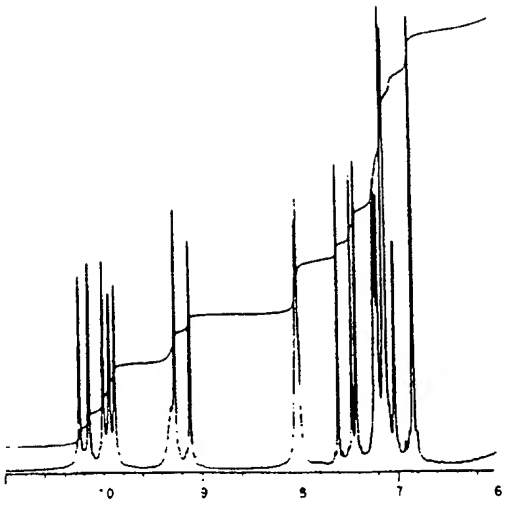


ImImPyPy-γ-ImOpPyPy-β-Dp

(2) ImImPyPy-γ-ImHpPyPy-β-Dp

FIG. 6

(A)



(B)

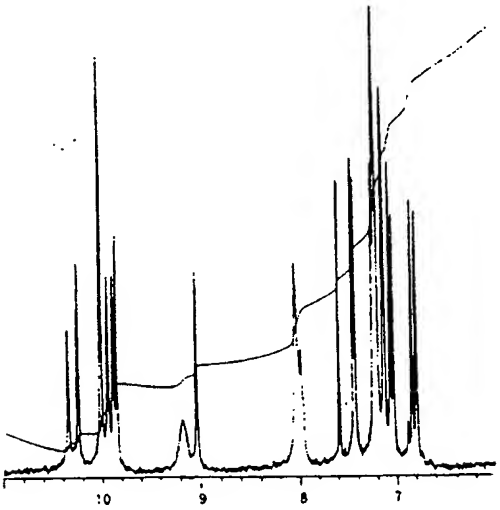
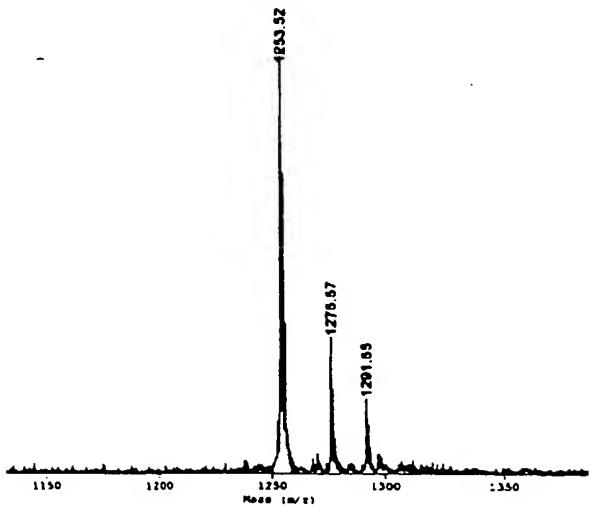


FIG. 7

(A)


1253.52
1278.57
1291.63
1150
1200
1250
1300
1350

(B)

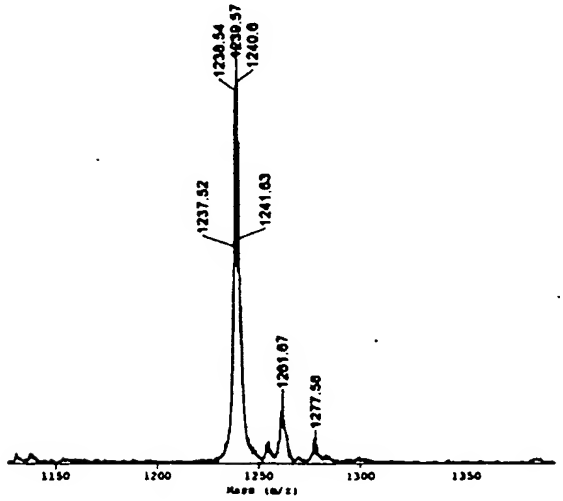
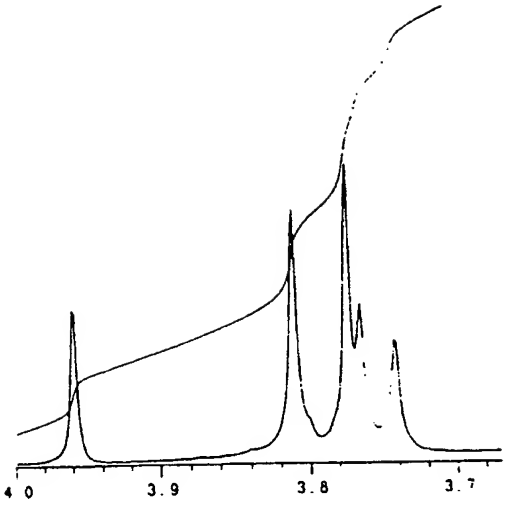

1238.54
1239.57
1240.6
1237.52
1241.63
1261.67
1277.58
1150
1200
1250
1300
1350

FIG. 8

(A)



(B)

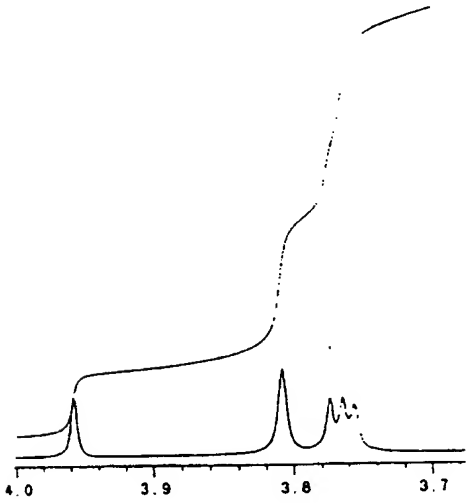


FIG. 9

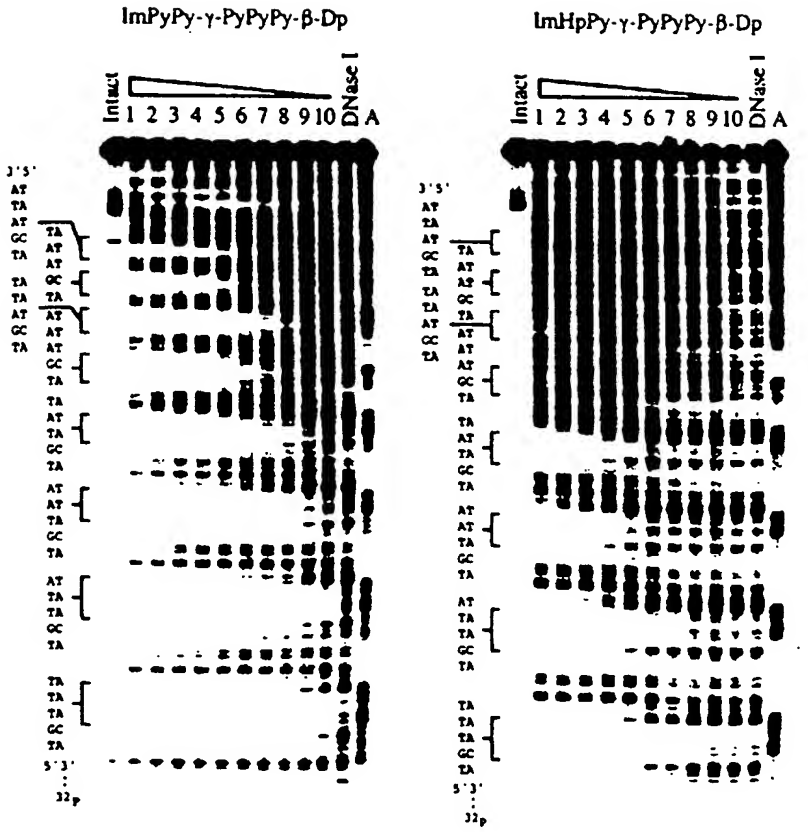
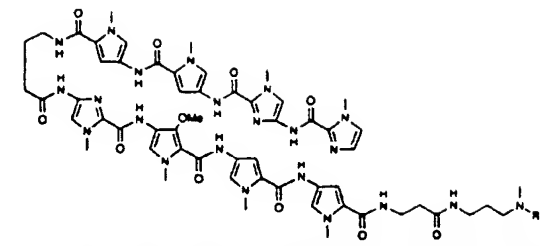
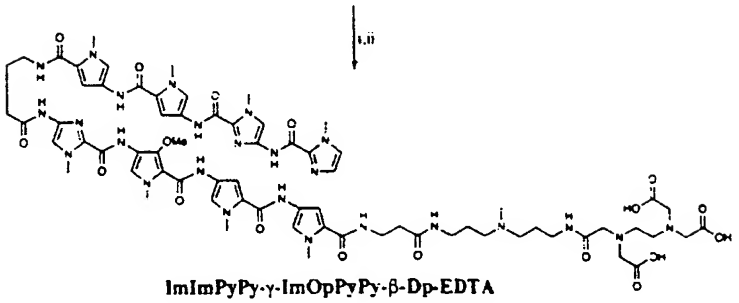

ImPyPy-γ-PyPyPy-β-Dp
ImHpPy-γ-PyPyPy-β-Dp
Intact
1 2 3 4 5 6 7 8 9 10
DNase I
A

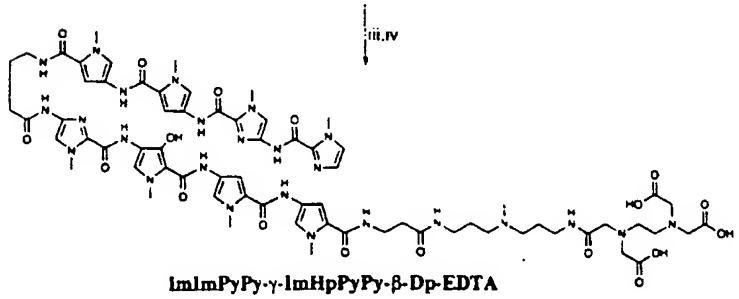
FIG. 10



R = $(CH_2)_3N(CH_3)_2$, ImImPyPy-γ-ImOpPyPy-β-Dp-NH_2



ImImPyPy-γ-ImOpPyPy-β-Dp-EDTA



ImImPyPy-γ-ImHpPyPy-β-Dp-EDTA

FIG. 11

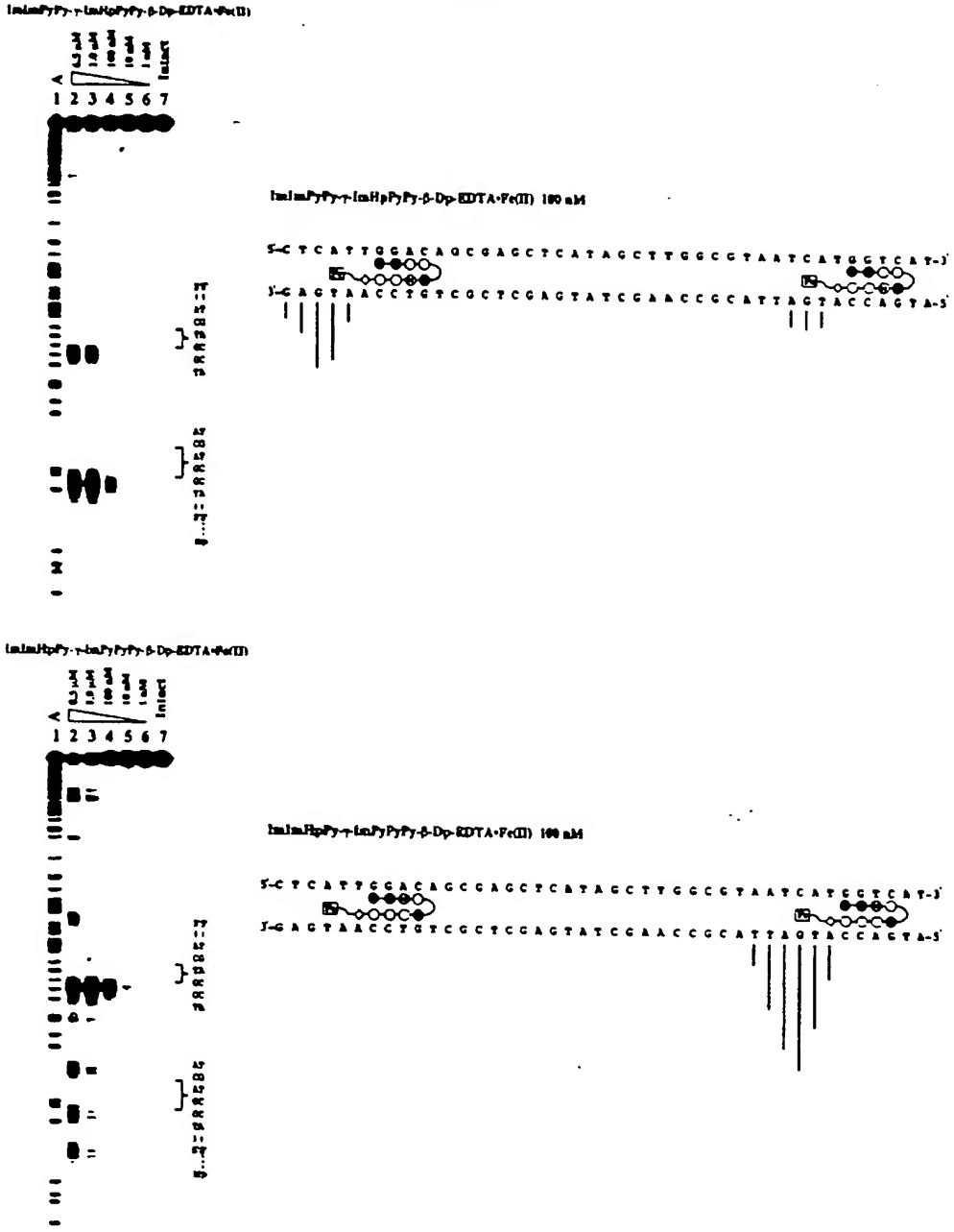


FIG. 12

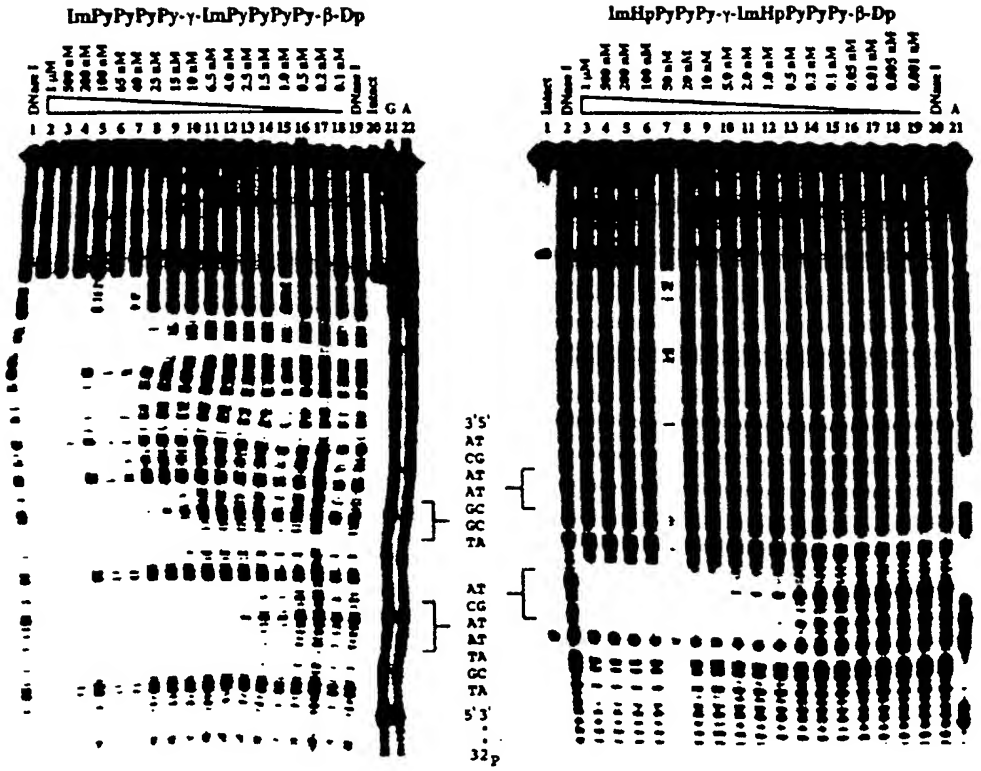

ImPyPyPyPy-γ-ImPyPyPyPy-β-Dp
ImHpPyPyPy-γ-ImHpPyPyPy-β-Dp
3'
5'
5'
3'
32p

FIG. 13

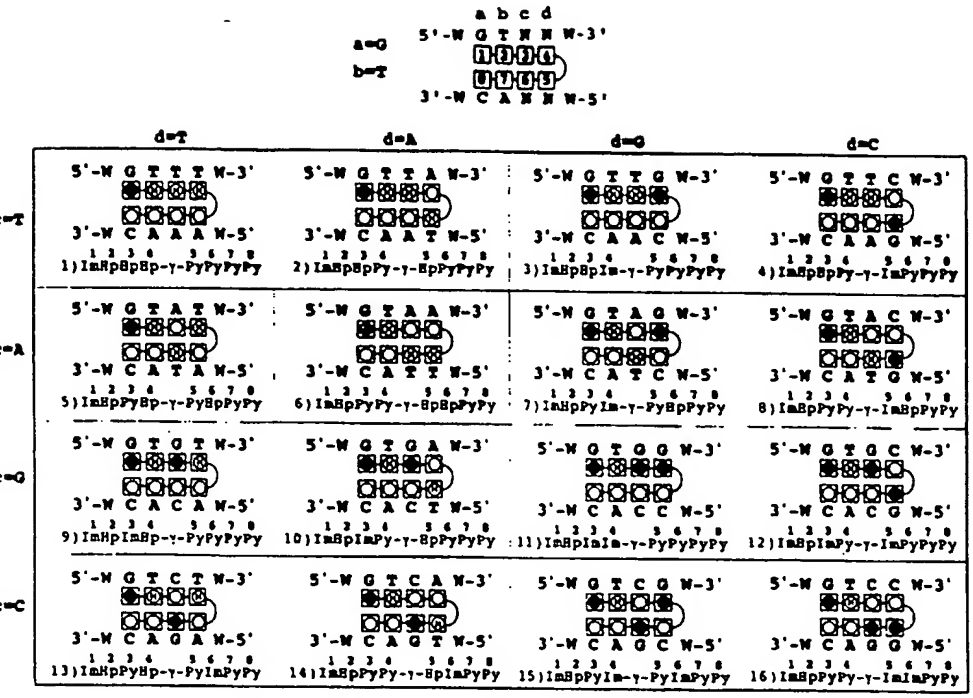

a b c d
a=G
b=T
5'-W G T N N W-3'
3'-W C A N N W-5'
d=T
d=A
d=G
d=C
c=T
c=A
c=G
c=C
1)ImHpBpBp-γ-PyPyPyPy
2)ImHpBpPy-γ-BpPyPyPy
3)ImHpBpIm-γ-PyPyPyPy
4)ImHpBpPy-γ-ImPyPyPy
5)ImBpPyBp-γ-PyBpPyPy
6)ImBpPyPy-γ-BpBpPyPy
7)ImBpPyIm-γ-PyBpPyPy
8)ImBpPyPy-γ-ImBpPyPy
9)ImHpImBp-γ-PyPyPyPy
10)ImHpImPy-γ-BpPyPyPy
11)ImHpImIm-γ-PyPyPyPy
12)ImHpImPy-γ-ImPyPyPy
13)ImHpPyHp-γ-PyImPyPy
14)ImHpPyPy-γ-HpImPyPy
15)ImHpPyIm-γ-PyImPyPy
16)ImHpPyPy-γ-ImImPyPy

FIG. 14

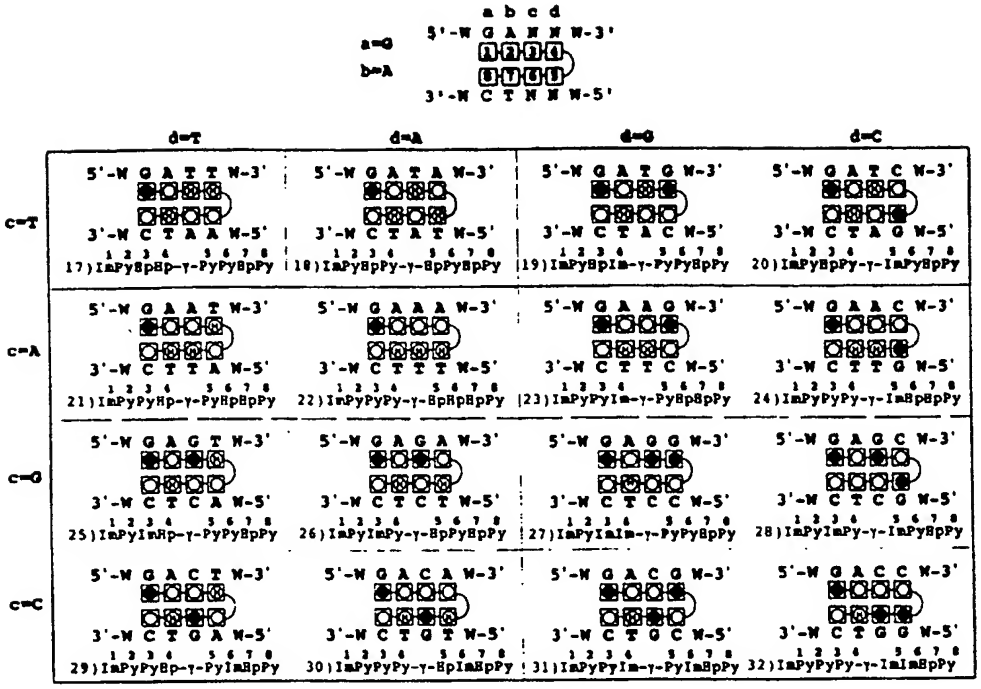

a b c d
5'-W G A N N W-3'
a=G
b=A
3'-W C T N N W-5'
d=T
d=A
d=G
d=C
c=T
c=A
c=G
c=C
5'-W G A T T W-3'
3'-W C T A A W-5'
17)ImPyBpHp-γ-PyPyBpPy
5'-W G A T A W-3'
3'-W C T A T W-5'
18)ImPyHpPy-γ-HpPyHpPy
5'-W G A T G W-3'
3'-W C T A C W-5'
19)ImPyHpIm-γ-PyPyHpPy
5'-W G A T C W-3'
3'-W C T A G W-5'
20)ImPyHpPy-γ-ImPyHpPy
5'-W G A A T W-3'
3'-W C T T A W-5'
21)ImPyPyHp-γ-PyHpHpPy
5'-W G A A A W-3'
3'-W C T T T W-5'
22)ImPyPyPy-γ-HpHpHpPy
5'-W G A A G W-3'
3'-W C T T C W-5'
23)ImPyPyIm-γ-PyHpHpPy
5'-W G A A C W-3'
3'-W C T T G W-5'
24)ImPyPyPy-γ-ImHpHpPy
5'-W G A G T W-3'
3'-W C T C A W-5'
25)ImPyImHp-γ-PyPyHpPy
5'-W G A G A W-3'
3'-W C T C T W-5'
26)ImPyImPy-γ-HpPyHpPy
5'-W G A G G W-3'
3'-W C T C C W-5'
27)ImPyImIm-γ-PyPyHpPy
5'-W G A G C W-3'
3'-W C T C G W-5'
28)ImPyImPy-γ-ImPyHpPy
5'-W G A C T W-3'
3'-W C T G A W-5'
29)ImPyPyHp-γ-PyImHpPy
5'-W G A C A W-3'
3'-W C T G T W-5'
30)ImPyPyPy-γ-HpImHpPy
5'-W G A C G W-3'
3'-W C T G C W-5'
31)ImPyPyIm-γ-PyImHpPy
5'-W G A C C W-3'
3'-W C T G G W-5'
32)ImPyPyPy-γ-ImImHpPy

FIG.15

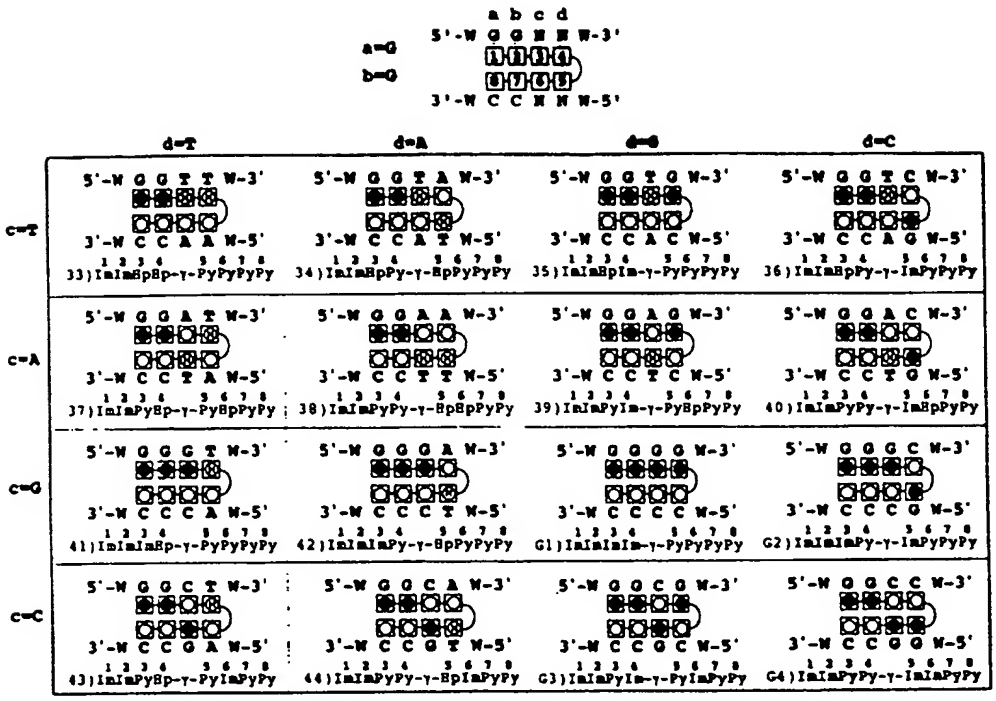

a b c d
5'-W G G N N W-3'
a=G
b=G
3'-W C C N N W-5'
d=T
d=A
d=G
d=C
c=T
c=A
c=G
c=C
33)ImImHpHp-γ-PyPyPyPy
34)ImImHpPy-γ-HpPyPyPy
35)ImImHpIm-γ-PyPyPyPy
36)ImImHpPy-γ-ImPyPyPy
37)ImImPyHp-γ-PyHpPyPy
38)ImImPyPy-γ-HpHpPyPy
39)ImImPyIm-γ-PyHpPyPy
40)ImImPyPy-γ-ImHpPyPy
41)ImImImHp-γ-PyPyPyPy
42)ImImImPy-γ-HpPyPyPy
G1)ImImImIm-γ-PyPyPyPy
G2)ImImImPy-γ-ImPyPyPy
43)ImImPyHp-γ-PyImPyPy
44)ImImPyPy-γ-HpImPyPy
G3)ImImPyIm-γ-PyImPyPy
G4)ImImPyPy-γ-ImImPyPy

FIG. 16

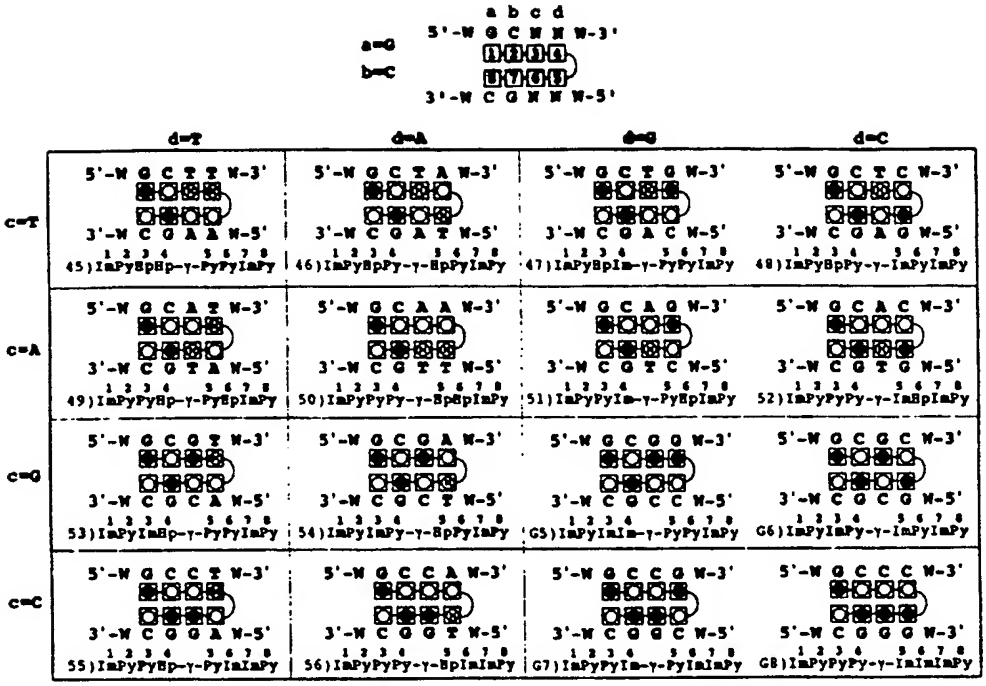

a b c d
5'-W G C W W W-3'
a=G
b=C
3'-W C G W W W-5'
d=T
d=A
d=G
d=C
c=T
c=A
c=G
c=C
45)ImPyBpHp-γ-PyPyImPy
46)ImPyBpPy-γ-HpPyImPy
47)ImPyBpIm-γ-PyPyImPy
48)ImPyBpPy-γ-ImPyImPy
49)ImPyPyHp-γ-PyBpImPy
50)ImPyPyPy-γ-BpBpImPy
51)ImPyPyIm-γ-PyBpImPy
52)ImPyPyPy-γ-ImBpImPy
53)ImPyImHp-γ-PyPyImPy
54)ImPyImPy-γ-BpPyImPy
G5)ImPyImIm-γ-PyPyImPy
G6)ImPyImPy-γ-ImPyImPy
55)ImPyPyHp-γ-PyImImPy
56)ImPyPyPy-γ-BpImImPy
G7)ImPyPyIm-γ-PyImImPy
G8)ImPyPyPy-γ-ImImImPy

FIG. 17

INTERNATIONAL SEARCH REPORT

Int. tional Application No
PCT/US 98/01006

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07D207/34 C07D233/90 A61K31/415 C07D403/14 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07D A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. W. TRAUGER ET AL: "Recognition of DNA by designed ligands at subnanomolar concentrations" NATURE, vol. 382, no. 6591, 8 August 1996, pages 559-561, XP002066256 cited in the application see the whole document ---	13-15,17
X	S.E.SWALLEY ET AL: "Recognition of a 5'-(A,T)GGG(A,T)2-3' sequence in the minor groove of DNA by an eight-ring hairpin polyamide" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 118, no. 35, 4 September 1996, pages 8198-8206, XP002066377 see page 8198 - page 8202 --- -/--	13-15,17

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 May 1998	12. 06. 98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Voyiazoglou, D

8

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/US 98/01006

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	E. B. BAIRD ET AL: "Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 118, no. 26, July 1996, pages 6141-6146, XP000674666 cited in the application see page 6141 - page 6142	13-15,17
A	M. E. PARKS ET AL: "Optimization of the hairpin polyamide design for recognition of the minor groove of DNA" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 118, no. 26, July 1996, pages 6147-6152, XP000674668 see page 6147 - page 6148	13-15,17
A	M. E. PARKS ET AL : "Recognition of 5'-(A,T)GG(AT)2-3' sequences in the minor groove of DNA by hairpin polyamides" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 118, no. 26, July 1996, DC US, pages 6153-6159, XP000674667 see page 6153 - page 6155	13-15,17
A	WO 96 05196 A (PHARMACIA) 22 February 1996 see claim 1	13-15,17
P,X	S. E. SWALLEY ET AL : "Discrimination of 5'-GGGG-3', and 5'-GGCC-3' sequences in the minor groove of DNA by eight-ring hairpin polyamides" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 119, no. 30, 30 July 1997, DC US, pages 6953-6961, XP002066260 see page 6959 - page 6961	13-15,17
P,X	W. L. WALKER ET AL: "Estimation of the DNA sequence discriminatory ability of hairpin-linked lexitropsins" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES,U.S.A., vol. 94, no. 11, May 1997, pages 5634-5639, XP002066261 see table 1	13-15,17

8

INTERNATIONAL SEARCH REPORT

Int tional application No.
PCT/US 98/01006

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 1-12,16
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

The claims are so broad that for determining the scope of a neaningful international search due account has been taken of rule 33.3 PCT; special emphasis was put on the following subject-matter: claims 13-15,17; pages 1-34; examples 1-8 and figures 1-11

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte ional Application No
PCT/US 98/01006

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9605196 A	22-02-1996	AU 689623 B	02-04-1998
		AU 3113695 A	07-03-1996
		CA 2172629 A	22-02-1996
		CN 1131946 A	25-09-1996
		EP 0722446 A	24-07-1996
		FI 961506 A	05-06-1996
		HU 76267 A	28-07-1997
		JP 9504039 T	22-04-1997
		NO 961377 A	30-05-1996
		NZ 290404 A	24-04-1997
		PL 313821 A	22-07-1996
		ZA 9506590 A	18-03-1996

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☑ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)